

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 12:53:09 ; Search time 62.9091 Seconds
(without alignments)
71.862 Million cell updates/sec

Title: US-09-336-091-3
Perfect score: 74
Sequence: 1 VKVLEYVIVKVSARVRF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 74 | 100.0 | 16 | 4 AAB31291 | Aab31291 Peptide d |
| 2 | 74 | 100.0 | 30 | 5 AAU85042 | Aau85042 Human MAG |
| 3 | 74 | 100.0 | 309 | 2 AAR70909 | Aar70909 Human mel |
| 4 | 74 | 100.0 | 309 | 2 AAW81548 | Aaw81548 Tumour re |
| 5 | 74 | 100.0 | 309 | 4 AAB31290 | Aab31290 Amino aci |
| 6 | 74 | 100.0 | 309 | 4 AAe06806 | Aae06806 Human MAG |
| 7 | 74 | 100.0 | 309 | 5 AAU84814 | Aau84814 Human MAG |
| 8 | 74 | 100.0 | 309 | 6 ABP74195 | Abp74195 Human MAG |
| 9 | 74 | 100.0 | 309 | 6 ABU08930 | Abu08930 Human tum |
| 10 | 74 | 100.0 | 309 | 7 ADC09573 | Adc09573 MAGE-1 pr |
| 11 | 74 | 100.0 | 310 | 6 AAO19742 | Aao19742 Wild-type |
| 12 | 74 | 100.0 | 316 | 6 ABU04419 | Abu04419 Human exp |
| 13 | 74 | 100.0 | 445 | 2 AAY06592 | Aay06592 CLYTA-MAG |
| 14 | 74 | 100.0 | 446 | 2 AAY06590 | Aay06590 Lipoprote |
| 15 | 74 | 100.0 | 1052 | 6 ABR57354 | Abr57354 MatDC16-C |
| 16 | 74 | 100.0 | 3541 | 5 AAU85130 | Aau85130 Human mel |
| 17 | 57 | 77.0 | 12 | 4 AAB31295 | Aab31295 Peptide d |
| 18 | 57 | 77.0 | 16 | 4 AAB31292 | Aab31292 Peptide d |
| 19 | 56 | 75.7 | 16 | 4 AAB31300 | Aab31300 Peptide d |
| 20 | 56 | 75.7 | 317 | 2 AAY06998 | Aay06998 MAGE-4 pr |
| 21 | 56 | 75.7 | 317 | 6 ABU56545 | Abu56545 Lung canc |
| 22 | 56 | 75.7 | 317 | 6 ADA83770 | Ada83770 Human MAG |
| 23 | 56 | 75.7 | 318 | 3 AAB08734 | Aab08734 Amino aci |
| 24 | 54 | 73.0 | 12 | 4 AAB31294 | Aab31294 Peptide d |
| 25 | 46 | 62.2 | 10 | 2 AAY38326 | Aay38326 MAGE-deri |

| | | | | | |
|----|----|------|----|------------|--------------------|
| 26 | 46 | 62.2 | 10 | 2 AAR70948 | Aar70948 Human mel |
| 27 | 46 | 62.2 | 10 | 2 AAR70957 | Aar70957 Human mel |
| 28 | 46 | 62.2 | 10 | 2 AAR70966 | Aar70966 Human mel |
| 29 | 46 | 62.2 | 10 | 2 AAR70960 | Aar70960 Human mel |
| 30 | 46 | 62.2 | 10 | 2 AAY46118 | Aay46118 Immunogen |
| 31 | 46 | 62.2 | 10 | 2 AAY46178 | Aay46178 Immunogen |
| 32 | 46 | 62.2 | 10 | 2 AAY45907 | Aay45907 Immunogen |
| 33 | 46 | 62.2 | 10 | 6 ABP74245 | Abp74245 Human MAG |
| 34 | 46 | 62.2 | 10 | 6 ABP74240 | Abp74240 Human MAG |
| 35 | 46 | 62.2 | 10 | 7 ADC09099 | Adc09099 Epitope w |
| 36 | 46 | 62.2 | 10 | 7 ADC09104 | Adc09104 Epitope w |
| 37 | 45 | 60.8 | 10 | 2 AAR70943 | Aar70943 Human mel |
| 38 | 45 | 60.8 | 10 | 2 AAR70953 | Aar70953 Human mel |
| 39 | 45 | 60.8 | 10 | 2 AAY47264 | Aay47264 Immunogen |
| 40 | 45 | 60.8 | 10 | 2 AAY46179 | Aay46179 Immunogen |
| 41 | 43 | 58.1 | 16 | 4 AAB31301 | Aab31301 Peptide d |
| 42 | 42 | 56.8 | 9 | 2 AAR73838 | Aar73838 Antigen f |
| 43 | 42 | 56.8 | 9 | 2 AAR78907 | Aar78907 MAGE 1 27 |
| 44 | 42 | 56.8 | 9 | 2 AAR70922 | Aar70922 Human mel |
| 45 | 42 | 56.8 | 9 | 2 AAY47408 | Aay47408 Immunogen |

ALIGNMENTS

RESULT 1
AAB31291
ID AAB31291 standard; peptide; 16 AA.
XX
AC AAB31291;
XX
DT 20-APR-2001 (first entry)
XX
DE Peptide derived from human MAGE-A1 HLA class II-binding protein.
XX
KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;
KW MAGE-A1 HLA class II-binding protein; vaccine.
XX
OS Homo sapiens.
XX
PN WO200078806-A1.
XX
PD 28-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US016287.
XX
PR 18-JUN-1999; 99US-00336091.
XX
(LUDW-) LUDWIG INST CANCER RES.
PI Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;
XX
DR WPI; 2001-102698/11.
XX
PT Novel MAGE-A1 human leukocyte antigen class II peptides which bind to and
PT are presented to the class II molecules, useful for inducing immune
PT response and treating cancers characterized by expression of MAGE-A1.
XX
PS Claim 3; Page 42; 78pp; English.
XX
CC The present sequence is derived from a human MAGE-A1 HLA (human leukocyte
CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA
CC binding protein stimulate the activity and proliferation of CD4+ T
CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic
CC agent for diagnosing a disorder characterized by expression of MAGE-A1.
CC The protein is used for treating a disorder characterized by expression
CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,
CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias. Peptides
CC derived from the MAGE-A1 HLA binding protein are useful in the production
XX of anti-tumour vaccines
SQ Sequence 16 AA;

Query Match 100.0%; Score 74; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | |
Db 1 VKVLEYVIKVSARVRF 16

RESULT 2
AAU85042
ID AAU85042 standard; peptide; 30 AA.
XX
AC AAU85042;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human MAGE-1 segment 19.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Homo sapiens.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
DR N-PSDB; ABK36862.
XX
PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.
XX
PS Example 3; Fig 27; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
CC a peptide derived from a parent protein used to construct a savine of the
CC invention
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 74; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | |
Db 9 VKVLEYVIKVSARVRF 24

RESULT 3
AAR70909
ID AAR70909 standard; protein; 309 AA.
XX
AC AAR70909;
XX
DT 25-MAR-2003 (revised)
DT 09-OCT-1995 (first entry)
XX
DE Human melanoma antigen MAGE-1.
XX
KW Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW HLA-restricted cytotoxic T-lymphocyte activity.
XX
OS Homo sapiens.
XX
PN WO9504542-A1.
XX
PD 16-FEB-1995.
XX
PF 02-AUG-1994; 94WO-US008721.
XX
PR 06-AUG-1993; 93US-00103623.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Fikes JD, Livingston BD, Sette AD, Sidney JC;
XX
DR WPI; 1995-090681/12.
DR N-PSDB; AAQ85435.
XX
PT Human melanoma antigen, MAGE-1, peptide(s) - useful for stimulating
PT immune response against melanoma.
XX
PS Example1; Fig 1; 59pp; English.
XX
CC AAQ85435 encodes AAR70909 human melanoma antigen MAGE-1, it was used to
CC produce the C-terminal MAGE-1 peptides described in AAR70915 to AAR70969.
CC These peptides are useful for defining epitopes that engender a HLA-
CC restricted cytotoxic lymphocyte activity against MAGE-1 antigens.
CC Compns. containing these peptides can be administered, as a vaccine to
CC patients susceptible to MAGE associated tumours, e.g. melanomas. (Updated
CC on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 74; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | |
Db 277 VKVLEYVIKVSARVRF 292

RESULT 4
AAW81548
ID AAW81548 standard; protein; 309 AA.
XX
AC AAW81548;
XX
DT 01-MAR-1999 (first entry)
XX

DE Tumour rejection antigen precursor MAGE-A1.
XX MAGE-A1; human; tumour rejection antigen precursor; TRAP; therapy;
KW diagnosis.
XX Homo sapiens.
OS
XX
PN WO9849184-A1.
XX
PD 05-NOV-1998.
XX
PF 24-APR-1998; 98WO-US008493.
XX
PR 25-APR-1997; 97US-00845528.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Lucas S, De Smet C, Boon-Falleur T;
XX
DR WPI; 1999-024041/02.
N-PSDB; AAV69719.
XX
PT Tumour rejection antigen precursors - used for determining presence of
PT cytolytic T cells specific for complexes of a human leukocyte antigen.
XX
PS Disclosure; Page 50-51; 84pp; English.
XX
CC This is the amino acid sequence of human tumour rejection antigen
CC precursor (TRAP) MAGE-A1. MAGE-A1 cDNA (see AAV69719) shows homology to
CC novel human MAGE-C1 cDNA (see AAV69720). MAGE-C1 (see AAW81546) is a
CC novel member of the MAGE family that may be recognised by cytotoxic T
CC cells, leading to lysis of the tumour cells which express it. It is
CC expressed in a variety of tumours and in normal testis cells, but not by
CC other normal cells. The invention provides MAGE-C1 and MAGE-C2 nucleic
CC acids and polypeptides, useful e.g. in a claimed method for determining
CC the presence of cytolytic T cells specific for complexes of a human
CC leukocyte antigen (HLA)
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 74; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
Db 277 VKVLEYVIKVSARVRF 292
|||||
RESULT 5
AAB31290
ID AAB31290 standard; protein; 309 AA.
XX
AC AAB31290;
XX
DT 20-APR-2001 (first entry)
XX
DE Amino acid sequence of human MAGE-A1 HLA class II-binding protein.
XX
KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;
KW MAGE-A1 HLA class II-binding protein; vaccine.
XX
OS Homo sapiens.
XX
PN WO200078806-A1.
XX
PD 28-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US016287.
XX
PR 18-JUN-1999; 99US-00336091.
XX
PA (LUDW-) LUDWIG INST CANCER RES.

XX Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;
PI WPI; 2001-102698/11.
XX N-PSDB; AAF24676.
DR
XX
PT Novel MAGE-A1 human leukocyte antigen class II peptides which bind to and
PT are presented to the class II molecules, useful for inducing immune
PT response and treating cancers characterized by expression of MAGE-A1.
XX
PS Claim 1; Page 63; 78pp; English.
XX
CC The present sequence represents a human MAGE-A1 HLA (human leukocyte
CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA
CC binding protein stimulate the activity and proliferation of CD4+ T
CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic
CC agent for diagnosing a disorder characterized by expression of MAGE-A1.
CC The protein is used for treating a disorder characterized by expression
CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,
CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias. Peptides
CC derived from the MAGE-A1 HLA binding protein are useful in the production
CC of anti-tumour vaccines
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 74; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
Db 277 VKVLEYVIKVSARVRF 292
|||||
RESULT 6
AAE06806
ID AAE06806 standard; protein; 309 AA.
XX
AC AAE06806;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human MAGE-A1 protein.
XX
KW MAGE antigenic peptide; Human leukocyte antigen; HLA-B35; HLA-B44;
KW tumour cell; immunostimulant; antigen presentation; cancer; melanoma;
KW CD8+ cytotoxic T lymphocyte; colorectal; prostate; gastric carcinoma;
KW myeloma; brain tumour; sarcoma; seminoma; ovarian tumour; cytostatic;
KW gene therapy; human; MAGE-A1; tumour rejection antigen; TRA.
XX
OS Homo sapiens.
XX
PN WO200153833-A1.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US002008.
XX
PR 20-JAN-2000; 2000US-0177242P.
PR 25-OCT-2000; 2000US-0243212P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Luiten R, Boon-Falleur T, Van Der Bruggen P, Stroobant V;
PI Demotte N, Schultz E;
XX
DR WPI; 2001-488724/53.
DR N-PSDB; AAD12987.
XX
PT Functional variants and isolated mimetics of a MAGE-A1 HLA-B35 or HLA-B44
PT binding peptide, or of a MAGE-A3 HLA-B35 binding peptide, used in
PT diagnosis and treatment of a disorder characterized by expression of MAGE
PT -A1 or -A3.

XX Claim 2; Page 86-87; 103pp; English.

PS The invention relates to functional variants and isolated mimetics of a

XX MAGE-A1 human leukocyte antigen (HLA)-B35 or HLA-B44 binding peptide, or

CC of a MAGE-A3 HLA-B35 binding peptide, identified by methods described in

CC the specification. MAGE genes encode tumour rejection antigens (TRAs)

CC presented to T lymphocytes by HLA-B35 and HLA-B44 molecules. The MAGE

CC antigenic peptide acts by binding to HLA molecules on tumour cells and

CC stimulating recognition of these cells and thus signalling them to the

CC immune system for destruction. The peptide when presented by HLA molecule

CC induces the activation and stimulation of CD8+ cytotoxic T lymphocytes.

CC The MAGE antigenic peptide is used to treat and diagnose disorders

CC characterised by expression of MAGE-A1 or -A3. Disorders include cancers

CC e.g melanomas, oesophageal, lung, head and neck, breast, colorectal,

CC prostate, renal, bladder, hepatocellular, papillary thyroid and gastric

CC carcinomas, myelomas, brain tumours, sarcomas, seminomas, and ovarian

CC tumours. The present sequence is human MAGE-A1 protein

XX

SQ Sequence 309 AA;

Query Match 100.0%; Score 74; DB 4; Length 309;

Best Local Similarity 100.0%; Pred. NO. 4.5e-05;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16

Db 277 VKVLEYVIKVSARVRF 292

RESULT 7

AAU84814

ID AAU84814 standard; protein; 309 AA.

XX

AC AAU84814;

XX

DT 08-MAY-2002 (first entry)

XX

DE Human MAGE-1 consensus sequence.

XX

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

KW viral infection; human immunodeficiency virus; melanoma;

KW bacterial infection; Salmonella; Legionella; parasitic infection;

KW Trypanosoma; Toxoplasma; Giardia.

XX

OS Homo sapiens.

XX

PN WO200190197-A1.

XX

PD 29-NOV-2001.

XX

PF 25-MAY-2001; 2001WO-AU000622.

XX

PR 26-MAY-2000; 2000AU-00007761.

XX

PA (AUSU) UNIV AUSTRALIAN NAT.

XX

PI Thomson SA, Ramshaw IA;

XX

DR WPI; 2002-147575/19.

XX

PT New synthetic polypeptides having several different segments of at least

PT one parent polypeptide linked together differently compared to the

PT linkage in the parent polypeptide, for inducing immune response against a

PT pathogen or cancer.

XX

PS Example 3; Fig 27; 364pp; English.

XX

CC The invention relates to a new synthetic polypeptide (I) comprising

CC several different segments of at least one parent polypeptide linked

CC together in a different relationship relative to their linkage in the

CC parent polypeptide to impede, abrogate or otherwise alter at least one

CC function associated with the parent polypeptide and for inducing an

CC immune response against a pathogen or cancer. Also included are a

CC synthetic polynucleotide encoding and a computer system for designing the

CC synthetic polypeptides. The synthetic polypeptides and polynucleotides

CC are referred to as a Savine. The synthetic polypeptide is useful for

CC modulating immune responses preferably directed against a pathogen or a

CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head

CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,

CC oesophagus, brain, testicle, uterus), as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or

CC prophylaxis against viral (such as infections caused by HIV (human

CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis

CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial

CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,

CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic

CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,

CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is

CC a consensus sequence for a parent protein used to design a savine of the

XX invention

SQ Sequence 309 AA;

Query Match 100.0%; Score 74; DB 5; Length 309;

Best Local Similarity 100.0%; Pred. No. 4.5e-05;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16

Db 277 VKVLEYVIKVSARVRF 292

RESULT 8

ABP74195

ID ABP74195 standard; protein; 309 AA.

XX

AC ABP74195;

XX

DT 03-FEB-2003 (first entry)

XX

DE Human MAGE-1 protein SEQ ID NO:71.

XX

KW Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;

KW T cell.

XX

OS Homo sapiens.

XX

PN WO200281646-A2.

XX

PD 17-OCT-2002.

XX

PF 04-APR-2002; 2002WO-US011101.

XX

PR 06-APR-2001; 2001US-0282211P.

PR 07-NOV-2001; 2001US-0337017P.

PR 07-MAR-2002; 2002US-0363210P.

XX

PA (CTLI-) CTL IMMUNOTHERAPIES CORP.

XX

PI Simard J JL, Diamond DC, Liu L, Xie Z;

XX

DR WPI; 2003-067518/06.

DR N-PSDB; ABQ83847.

XX

PT Novel epitopes useful as vaccines, comprises peptides or nucleic acid

PT encoding the peptides, that are useful epitopes of target-associated

PT antigens.

XX

PS Claim 1; Page 156; 352pp; English.

XX

CC The present invention describes an isolated epitope (I) and an epitope

CC cluster. Also described is a vaccine or immunotherapeutic composition

CC (VC) comprising (I). (I) has cytostatic activity. VC is useful for

CC treating an animal, by administering to an animal the vaccine or

CC immunotherapeutic composition. VC is also useful for evaluating

CC immunogenicity of a vaccine or immunotherapeutic composition, by
CC administering VC to an HLA-transgenic animal and evaluating
CC immunogenicity based on a characteristic of the animal, or by in vitro
CC primary stimulation of a T cell and evaluating immunogenicity. (I) is
CC useful for determining specific T cell frequency, by contacting T cells
CC with a MHC-peptide complex, and further comprises ELISPOT analysis,
CC limiting dilution analysis, flow cytometry, in situ hybridisation and/or
CC polymerase chain reaction (PCR). ABQ83843 to ABQ83858 and ABP74128 to
CC ABP74713 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 74; DB 6; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | | | |
Db 277 VKVLEYVIKVSARVRF 292

RESULT 9
ABU08930
ID ABU08930 standard; protein; 309 AA.
XX
AC ABU08930;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human tumour rejection antigen precursor, MAGE-A1.
XX
KW TRAP; tumour rejection antigen precursor; cytolytic T-cell; CTL; tumour;
KW seminoma; bladder transitional-cell carcinoma; NSCLC; adaptor;
KW head-and-neck squamous-cell carcinoma; breast carcinoma; sarcoma;
KW cutaneous melanoma; non-small cell lung cancer; MAGE-A1; human.
XX
OS Homo sapiens.
XX
PN US2002176865-A1.
XX
PD 28-NOV-2002.
XX
PF 01-MAR-2002; 2002US-00085108.
XX
PR 25-APR-1997; 97US-00845528.
PR 24-APR-1998; 98US-00066281.
PR 17-DEC-1999; 99US-00468433.
PR 09-FEB-2000; 2000US-00501104.
XX
PA (LUCA/) LUCAS S.
PA (BOON/) BOON-FALLEUR T.
XX
PI Lucas S, Boon-Falleur T;
XX
DR WPI; 2003-328468/31.
DR N-PSDB; ABX93696.
XX
PT Novel isolated nucleic acid encoding tumor rejection antigen precursor
PT MAGE-C3, MAGE-B5, or MAGE-B6, useful as diagnostic probes to determine
PT presence of abnormal e.g., tumor cells expressing MAGE-C1, MAGE-B5 or
PT MAGE-B6.
XX
PS Disclosure; Fig 2; 59pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule which encodes
CC a tumour rejection antigen precursor (TRAP) having an amino acid sequence
CC of a TRAP encoded by a fully defined MAGE-C3, MAGE-B5, or MAGE-B6
CC polynucleotide sequence. Also disclosed is a method which is useful for
CC determining presence of cytolytic T-cells specific for complexes of human
CC leukocyte antigen (HLA) and a peptide derived from the nucleic acid in a
CC cytotoxic T-lymphocyte (CTL)-containing sample. The nucleic acid is
CC useful as a diagnostic probe to determine the presence of abnormal

CC (tumour) cells such as seminoma, bladder transitional-cell carcinoma,
CC head-and-neck squamous-cell carcinoma, breast carcinoma, sarcoma,
CC cutaneous melanoma or non-small cell lung cancer (NSCLC) which express
CC MAGE-C1, MAGE-B5 or MAGE-B6. The nucleic acid is useful for diagnosing a
CC disorder characterised by expression of MAGE-C1, MAGE-B5 or MAGE-B6 TRAPs
CC or tumour rejection antigens (TRAs). The present sequence represents the
CC amino acid sequence of the human tumour rejection antigen precursor, MAGE
CC -A1
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 74; DB 6; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | | | |
Db 277 VKVLEYVIKVSARVRF 292

RESULT 10
ADC09573
ID ADC09573 standard; protein; 309 AA.
XX
AC ADC09573;
XX
DT 18-DEC-2003 (first entry)
XX
DE MAGE-1 protein #SEQ ID 71.
XX
KW Epitope; immunological; vaccine;
KW major histocompatibility complex class I; MHC class I; cancer;
KW immunisation.
XX
OS Unidentified.
XX
PN WO2003008537-A2.
XX
PD 30-JAN-2003.
XX
PF 29-MAR-2002; 2002WO-US010189.
XX
PR 06-APR-2001; 2001US-0282211P.
PR 07-NOV-2001; 2001US-0337017P.
PR 07-MAR-2002; 2002US-0363210P.
XX
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX
PI Simard J.J.L., Diamond DC, Liu L, Xie Z;
XX
DR WPI; 2003-248010/24.
XX
PT Epitope having high affinity for major histocompatibility complex class I
PT useful for treating an animal, evaluating immunogenicity of a vaccine or
PT therapeutic composition and for diagnosing a disease.
XX
PS Claim 1; SEQ ID NO 71; 239pp; English.
XX
CC The invention relates to an isolated epitope polypeptide that has high
CC affinity for major histocompatibility complex (MHC) class I, and an
CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine
CC or immunotherapeutic composition containing an epitope of the invention.
CC Compositions of the invention may be used in the treatment of cancer. The
CC method can be combined with a radiation therapy, chemotherapy,
CC biochemotherapy or surgery. The composition is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
CC-peptide complexes of the invention are useful for determining specific T
CC cell frequency. This method is useful for evaluating immunological
CC response, by performing the method prior to and subsequent to an
CC immunisation step. Compositions of the invention are useful for
CC diagnosing a disease. The current sequence represents an epitope of the
CC invention with high affinity for MHC class I.
XX

SQ Sequence 309 AA;
Query Match 100.0%; Score 74; DB 7; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | | | |
Db 277 VKVLEYVIKVSARVRF 292
RESULT 11
AAO19742
ID AAO19742 standard; protein; 310 AA.
XX
AC AAO19742;
XX
DT 11-AUG-2003 (first entry)
XX
DE Wild-type MAGE1 protein.
XX
KW Stabilised mRNA; translation optimised; vaccine; tissue repair;
KW sequence modification determination; gene therapy; cytostatic; virucide;
KW antibacterial; protozoacide; nootropic; neuroprotective; infection;
KW antiparkinsonian; immunostimulant; cancer; MAGE1 protein.
XX
OS Unidentified.
XX
PN WO200298443-A2.
XX
PD 12-DEC-2002.
XX
PF 05-JUN-2002; 2002WO-EP006180.
XX
PR 05-JUN-2001; 2001DE-01027283.
XX
PA (VMUE/) VON DER MUELBE F.
XX
PI Von Der Muelbe F, Hoerr I, Pascolo S;
XX
DR WPI; 2003-148621/14.
DR N-PSDB; ABZ69107.
XX
PT Composition containing mRNA modified for optimal translation and
PT stability, useful for treating e.g. tumors or infections, comprises
PT increased G/C content and fewer rare codons.
XX
PS Disclosure; Fig 2B; 75pp; German.
XX
CC The present invention relates to a pharmaceutical composition containing
CC at least one modified RNA encoding a biologically active or antigenic
CC protein. The RNA is modified to optimise translation of the sequence. The
CC compositions are used for vaccination against a wide range of infectious
CC diseases (viral, bacterial or protozoal) or cancer, or for tissue
CC regeneration, e.g. in cases of Alzheimer's or Parkinson's diseases and
CC arthritis, but also to express proteins such as dystrophins, chloride ion
CC channels (for treating cystic fibrosis) and enzymes (either for treating
CC metabolic disorders or for synthesis of neurotransmitters such as
CC dopamine). The present sequence is the wild-type MAGE1 protein
XX
SQ Sequence 310 AA;
Query Match 100.0%; Score 74; DB 6; Length 310;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | | | |
Db 276 VKVLEYVIKVSARVRF 291
RESULT 12
ABU04419

ID ABU04419 standard; protein; 316 AA.
XX
AC ABU04419;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1085.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicz RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 1085; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 316 AA;
Query Match 100.0%; Score 74; DB 6; Length 316;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | | | |
Db 284 VKVLEYVIKVSARVRF 299
RESULT 13
AAU06592

ID AAY06592 standard; protein; 445 AA.
XX
AC AAY06592;
XX
DT 26-OCT-1999 (first entry)
XX
DE CLYTA-MAGE-1-His fusion protein.
XX
XX
KW MAGE-1; CLYTA-MAGE-1-His; fusion protein; tumour; melanoma;
KW breast cancer; bladder cancer; lung cancer; colon cancer;
KW head and squamous cell carcinoma; oesophagus carcinoma; vaccine; human.
XX
OS Streptococcus pneumoniae.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
PN WO9940188-A2.
XX
PD 12-AUG-1999.
XX
XX 02-FEB-1999; 99WO-EP0000660.
PF
XX 05-FEB-1998; 98GB-00002543.
PR
XX 06-FEB-1998; 98GB-00002650.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cabezon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;
XX
DR WPI; 1999-494293/41.
DR N-PSDB; AAX87593.
XX
PT New protein derivatives used in cancer vaccine therapy for treating a
range of cancers including melanomas, carcinomas and cancers of breast.
XX
PS Example 9; Page 69-70; 72pp; English.
XX
CC The present sequence represents a fusion protein composed of the C-
terminal portion of the Streptococcus pneumoniae LYTA protein (CLYTA),
the human MAGE-1 tumour-associated antigen and a hexahistidine tail. A
vector designed for recombinant expression of the fusion protein in
Escherichia coli is provided. The CLYTA moiety provides expression of
soluble fusion protein, facilitates affinity purification, and also acts
as a T-helper epitope. The invention relates to MAGE proteins fused to an
immunological fusion partner, e.g. CLYTA-MAGE-1-His. These novel fusion
proteins provide vaccines for immunotherapy of melanomas or other MAGE-
associated tumours like breast, bladder, lung and non-small cell lung
cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus
carcinoma
XX
SQ Sequence 445 AA;

Query Match 100.0%; Score 74; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
DB 404 VKVLEYVIKVSARVRF 419

RESULT 14
AAY06590
ID AAY06590 standard; protein; 446 AA.
XX
AC AAY06590;
XX
DT 26-OCT-1999 (first entry)
XX
DE Lipoprotein D-MAGE-1-His fusion protein.
XX
KW MAGE-1; lipoprotein D; LPD-MAGE-1-His; fusion protein; tumour; melanoma;

KW breast cancer; bladder cancer; lung cancer;
KW head and squamous cell carcinoma; colon cancer; oesophagus carcinoma;
XX vaccine; human.
XX
OS Haemophilus influenzae.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
PN WO9940188-A2.
XX
PD 12-AUG-1999.
XX
XX 02-FEB-1999; 99WO-EP0000660.
PF
XX 05-FEB-1998; 98GB-00002543.
PR
XX 06-FEB-1998; 98GB-00002650.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cabezon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;
XX
DR WPI; 1999-494293/41.
DR N-PSDB; AAX87591.
XX
PT New protein derivatives used in cancer vaccine therapy for treating a
range of cancers including melanomas, carcinomas and cancers of breast.
XX
PS Example 6; Page 67-68; 72pp; English.
XX
CC The present sequence represents a novel fusion protein composed of
lipidated protein D (LPD) of Haemophilus influenzae B, the human MAGE-1
tumour-associated antigen and a hexahistidine tail. The invention relates
to MAGE proteins fused to an immunological fusion partner such as LPD.
The LPD moiety provides the fusion protein with additional exogenous T-
cell epitopes and also increase expression levels in E. coli. The lipid
tail ensures optimal presentation of the antigen to antigen-presenting
cells. The affinity tag facilitates purification. The novel fusion
proteins provide vaccines for immunotherapy of melanomas or other MAGE-
associated tumours like breast, bladder, lung and non-small cell lung
cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus
carcinoma
XX
SQ Sequence 446 AA;

Query Match 100.0%; Score 74; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
DB 405 VKVLEYVIKVSARVRF 420

RESULT 15
ABR57354
ID ABR57354 standard; protein; 1052 AA.
XX
AC ABR57354;
XX
DT 09-SEP-2003 (first entry)
XX
DE MatDC16'-C-gamma-4-MAGE-A1 amino acid sequence.
XX
KW Antigen presenting cell; vaccination; nootropic; neuroprotective;
KW antiarteriosclerotic; cytostatic; antidiabetic; hepatotropic;
KW antiinflammatory; antiparasitic; fungicide; antibacterial; virucide;
KW vaccine; Alzheimer's disease; atherosclerosis; cancer; diabetes;
KW hepatitis; infection.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers

FT Misc-difference 546 /note= "unspecified"

XX WO2003046011-A1.

XX PD 05-JUN-2003.

XX PF 30-NOV-2001; 2001WO-EP014255.

XX PR 30-NOV-2001; 2001WO-EP014255.

XX PA (CRUC-) CRUCCELL HOLLAND BV.

XX PI Germeraad W;

XX WPI; 2003-493401/46.

PT New conjugate for targeting antigen presenting cells, useful for preventing, retarding or treating e.g., Alzheimer's disease, atherosclerosis, cancer, diabetes, hepatitis or fungal, bacterial or viral infections.

PS Disclosure; Fig 2; 54pp; English.

XX The present invention describes a conjugate (I) for targeting antigen presenting cells (APCs) comprising at least one antigenic moiety conjugated to a targeting moiety that is capable of binding to a cell surface structure of an APC, and upon binding, inducing a cytotoxic T lymphocyte (CTL) and T-helper response. Also described: (1) a nucleic acid sequence encoding the antigenic or targeting moiety; (2) an expression vector comprising the nucleic acid sequence, operably linked to expression sequences for the APC; (3) a host cell transformed or transfected using the nucleic acid or expression vector; (4) a method for producing (I); (5) a method for generating an APC, capable of eliciting an immune response via MHC classes I and II presentation of processed antigen fragments; and (6) a pharmaceutical composition comprising (I) or the APC. (I) has nootropic, neuroprotective, virucide, antiarteriosclerotic, cytostatic, antidiabetic, hepatotropic, fungicide, antiinflammatory, antiparasitic and antibacterial activities, and can be used in vaccines. The conjugate (I) or APC can be used for preventing, retarding or treating e.g., Alzheimer's disease, atherosclerosis, cancer, diabetes, hepatitis or parasitic, fungal, bacterial or viral infections. The present sequence represents a MatDC16-C-gamma-4-MAGE-A1 amino acid sequence, which is used in the exemplification of the present invention

XX SQ Sequence 1052 AA;

Query Match 100.0%; Score 74; DB 6; Length 1052;
Best Local Similarity 100.0%; Pred. NO. 0.00017;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRP 16

Db 1020 VKVLEYVIKVSARVRP 1035

Search completed: October 7, 2004, 15:13:01
Job time : 64.9091 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:30:52 ; Search time 59.6364 Seconds
(without alignments)
86.336 Million cell updates/sec

Title: US-09-336-091-3
Perfect score: 74
Sequence: 1 VKVLEYVIKVSARVRF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------------------------|
| 1 | 74 | 100.0 | 30 | 12 | US-10-296-734-1278 Sequence 1278, Ap |
| 2 | 74 | 100.0 | 309 | 9 | US-09-766-889A-2 Sequence 2, Appli |
| 3 | 74 | 100.0 | 309 | 12 | US-10-296-734-828 Sequence 828, App |
| 4 | 74 | 100.0 | 309 | 13 | US-10-085-108-10 Sequence 10, Appl |
| 5 | 74 | 100.0 | 309 | 14 | US-10-177-390-18 Sequence 18, Appl |
| 6 | 74 | 100.0 | 309 | 14 | US-10-160-237-10 Sequence 10, Appl |
| 7 | 74 | 100.0 | 309 | 15 | US-10-117-937-71 Sequence 71, Appl |
| 8 | 74 | 100.0 | 309 | 16 | US-10-657-022-71 Sequence 71, Appl |
| 9 | 74 | 100.0 | 309 | 16 | US-10-741-466-6 Sequence 6, Appli |
| 10 | 74 | 100.0 | 311 | 16 | US-10-741-466-7 Sequence 7, Appli |
| 11 | 74 | 100.0 | 311 | 16 | US-10-741-466-8 Sequence 8, Appli |
| 12 | 74 | 100.0 | 3541 | 12 | US-10-296-734-1454 Sequence 1454, Ap |
| 13 | 57 | 77.0 | 12 | 15 | US-10-164-121A-33 Sequence 33, Appl |
| 14 | 57 | 77.0 | 12 | 15 | US-10-164-078A-32 Sequence 32, Appl |
| 15 | 56 | 75.7 | 317 | 12 | US-10-218-095-2 Sequence 2, Appli |

| | | | | | | |
|----|----|------|------|----|----------------------|-------------------|
| 16 | 56 | 75.7 | 317 | 14 | US-10-157-031-52 | Sequence 52, Appl |
| 17 | 46 | 62.2 | 10 | 15 | US-10-117-937-124 | Sequence 124, App |
| 18 | 46 | 62.2 | 10 | 15 | US-10-117-937-129 | Sequence 129, App |
| 19 | 42 | 56.8 | 9 | 14 | US-10-128-711-84 | Sequence 84, Appl |
| 20 | 42 | 56.8 | 9 | 14 | US-10-128-711-143 | Sequence 143, App |
| 21 | 42 | 56.8 | 9 | 14 | US-10-150-797-1 | Sequence 1, Appli |
| 22 | 42 | 56.8 | 9 | 15 | US-10-117-937-113 | Sequence 113, App |
| 23 | 42 | 56.8 | 9 | 15 | US-10-117-937-123 | Sequence 123, App |
| 24 | 41 | 55.4 | 9 | 12 | US-10-149-135-2182 | Sequence 2182, Ap |
| 25 | 41 | 55.4 | 9 | 12 | US-09-935-476-6 | Sequence 6, Appli |
| 26 | 41 | 55.4 | 9 | 14 | US-10-128-711-93 | Sequence 93, Appl |
| 27 | 41 | 55.4 | 9 | 14 | US-10-128-711-152 | Sequence 152, App |
| 28 | 41 | 55.4 | 9 | 14 | US-10-150-797-13 | Sequence 13, Appl |
| 29 | 41 | 55.4 | 9 | 15 | US-10-149-138-4195 | Sequence 4195, Ap |
| 30 | 41 | 55.4 | 9 | 16 | US-10-149-138-4195 | Sequence 4195, Ap |
| 31 | 41 | 55.4 | 28 | 12 | US-10-296-734-1280 | Sequence 1280, Ap |
| 32 | 41 | 55.4 | 592 | 12 | US-10-282-122A-44745 | Sequence 44745, A |
| 33 | 40 | 54.1 | 319 | 15 | US-10-369-493-10934 | Sequence 10934, A |
| 34 | 40 | 54.1 | 823 | 16 | US-10-437-963-145378 | Sequence 145378, |
| 35 | 40 | 54.1 | 984 | 16 | US-10-437-963-105777 | Sequence 105777, |
| 36 | 39 | 52.7 | 9 | 15 | US-10-117-937-128 | Sequence 128, App |
| 37 | 38 | 51.4 | 218 | 16 | US-10-437-963-201197 | Sequence 201197, |
| 38 | 38 | 51.4 | 295 | 12 | US-10-335-977-6929 | Sequence 6929, Ap |
| 39 | 38 | 51.4 | 295 | 12 | US-10-335-977-6930 | Sequence 6930, Ap |
| 40 | 38 | 51.4 | 345 | 12 | US-10-424-599-186286 | Sequence 186286, |
| 41 | 38 | 51.4 | 487 | 16 | US-10-203-927A-8 | Sequence 8, Appli |
| 42 | 38 | 51.4 | 597 | 12 | US-10-282-122A-63003 | Sequence 63003, A |
| 43 | 38 | 51.4 | 761 | 15 | US-10-369-493-129 | Sequence 129, App |
| 44 | 38 | 51.4 | 846 | 16 | US-10-203-927A-12 | Sequence 12, Appl |
| 45 | 38 | 51.4 | 1034 | 16 | US-10-437-963-201195 | Sequence 201195, |

ALIGNMENTS

RESULT 1
US-10-296-734-1278
; Sequence 1278, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1278
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MAGE-1 segment 19
US-10-296-734-1278

Query Match 100.0%; Score 74; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | | | |
Db 9 VKVLEYVIKVSARVRF 24

RESULT 2
US-09-766-889A-2
; Sequence 2, Application US/09766889A
; Patent No. US20020164654A1
; GENERAL INFORMATION:
; APPLICANT: Luiten, Rosalie

```

; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: van der Bruggen, Pierre
; APPLICANT: Stroobant, Vincent
; APPLICANT: Demotte, Nathalie
; APPLICANT: Schultz, Erwin
; TITLE OF INVENTION: MAGE ANTIGENIC PEPTIDES WHICH BIND HLA-B35 AND HLA-B44
; FILE REFERENCE: L0461/7104
; CURRENT APPLICATION NUMBER: US/09/766,889A
; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/177,242
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/243,212
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-766-889A-2

Query Match      100.0%; Score 74; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VKVLEYVIKVSARVRF 16
Db      277 VKVLEYVIKVSARVRF 292

RESULT 3
US-10-296-734-828
; Sequence 828, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 828
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MAGE-1 consensus polypeptide
US-10-296-734-828

Query Match      100.0%; Score 74; DB 12; Length 309;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VKVLEYVIKVSARVRF 16
Db      277 VKVLEYVIKVSARVRF 292

RESULT 4
US-10-085-108-10
; Sequence 10, Application US/10085108
; Publication No. US20020176865A1
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING
; FOR
; TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C
; MAGE-B FAMILIES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/085,108
; FILING DATE: 01-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/501,104
; FILING DATE: 09-Feb-2000
; APPLICATION NUMBER: 09/468,433
; FILING DATE: December 17, 1999
; APPLICATION NUMBER: 09/066,281
; FILING DATE: April 24, 1998
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5611.1 JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 318-3400
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-085-108-10

Query Match      100.0%; Score 74; DB 13; Length 309;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VKVLEYVIKVSARVRF 16
Db      277 VKVLEYVIKVSARVRF 292

RESULT 5
US-10-177-390-18
; Sequence 18, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerold
; APPLICANT: N.V. Antwerps Innovatiecentrum
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with Linear
; FILE REFERENCE: 021505wo/JH/ml
; CURRENT APPLICATION NUMBER: US/10/177,390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-18

Query Match      100.0%; Score 74; DB 14; Length 309;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Qy 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | |
Db 277 VKVLEYVIKVSARVRF 292

RESULT 6
US-10-160-237-10
; Sequence 10, Application US/10160237
; Publication No. US20030170256A1
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; DE SMET, Charles; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING
; FOR TUMOR REJECTION ANTIGEN PRECURSOR MAGE-C1 AND MAGE-C2
; AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/160,237
; FILING DATE: 04-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,281B
; FILING DATE: April 24, 1998
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455.2 US - JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-160-237-10

Query Match 100.0%; Score 74; DB 14; Length 309;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | |
Db 277 VKVLEYVIKVSARVRF 292

RESULT 7
US-10-117-937-71
; Sequence 71, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES

; FILE REFERENCE: CTLIMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-117-937-71

Query Match 100.0%; Score 74; DB 15; Length 309;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | |
Db 277 VKVLEYVIKVSARVRF 292

RESULT 8
US-10-657-022-71
; Sequence 71, Application US/10657022
; Publication No. US20040180354A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Liu, Liping
; APPLICANT: Liu, Zheng
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: MANNK.032A
; CURRENT APPLICATION NUMBER: US/10/657,022
; CURRENT FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: 60/409123
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 610
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-657-022-71

Query Match 100.0%; Score 74; DB 16; Length 309;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | |
Db 277 VKVLEYVIKVSARVRF 292

RESULT 9
US-10-741-466-6
; Sequence 6, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1

; APPLICANT: Boon, Thierry
; TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-B18 Molecules And Uses Thereof
; FILE REFERENCE: LUD-5756
; CURRENT APPLICATION NUMBER: US/10/164,078A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 32
; LENGTH: 12
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
US-10-164-078A-32

Query Match 77.0%; Score 57; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EYIKVSARVF 16
| | | | | | | | | |
Db 1 EYIKVSARVF 12

RESULT 15
US-10-218-095-2
; Sequence 2, Application US/10218095
; Publication No. US20040033541A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Yi
; APPLICANT: Stroobant, Vincent
; APPLICANT: Russo, Vincenzo
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: van der Bruggen, Pierre
; TITLE OF INVENTION: MAGE-A4 ANTIGENIC PEPTIDES AND USES THEREOF
; FILE REFERENCE: L00461/70137
; CURRENT APPLICATION NUMBER: US/10/218,095
; CURRENT FILING DATE: 2002-08-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-095-2

Query Match 75.7%; Score 56; DB 12; Length 317;
Best Local Similarity 73.3%; Pred. No. 0.12;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKVLEYVIVKSARVR 15
| | | | | : : : : :
Db 285 VKVLEHVVRNARVR 299

Search completed: October 7, 2004, 15:33:46
Job time : 60.6364 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:18:26 ; Search time 17.4545 Seconds
(without alignments)
47.324 Million cell updates/sec

Title: US-09-336-091-3
Perfect score: 74
Sequence: 1 VKVLEYVIKVSARVRF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 74 | 100.0 | 58 | 1 US-08-465-167A-1 | Sequence 1, Appli |
| 2 | 74 | 100.0 | 58 | 4 US-08-627-820-1 | Sequence 1, Appli |
| 3 | 74 | 100.0 | 309 | 1 US-08-465-167A-24 | Sequence 24, Appl |
| 4 | 74 | 100.0 | 309 | 2 US-08-993-118-10 | Sequence 10, Appl |
| 5 | 74 | 100.0 | 309 | 3 US-08-845-528C-10 | Sequence 10, Appl |
| 6 | 74 | 100.0 | 309 | 4 US-08-627-820-24 | Sequence 24, Appl |
| 7 | 74 | 100.0 | 309 | 4 US-09-066-281B-10 | Sequence 10, Appl |
| 8 | 74 | 100.0 | 309 | 4 US-09-468-433C-10 | Sequence 10, Appl |
| 9 | 74 | 100.0 | 309 | 4 US-09-392-714-29 | Sequence 29, Appl |
| 10 | 46 | 62.2 | 10 | 1 US-08-465-167A-15 | Sequence 15, Appl |
| 11 | 46 | 62.2 | 10 | 1 US-08-465-167A-49 | Sequence 49, Appl |
| 12 | 46 | 62.2 | 10 | 3 US-08-159-339A-595 | Sequence 595, App |
| 13 | 46 | 62.2 | 10 | 4 US-08-627-820-15 | Sequence 15, Appl |
| 14 | 45 | 60.8 | 10 | 1 US-08-465-167A-39 | Sequence 39, Appl |
| 15 | 42 | 56.8 | 9 | 1 US-08-465-167A-4 | Sequence 4, Appli |
| 16 | 42 | 56.8 | 9 | 4 US-08-197-484-84 | Sequence 84, Appl |
| 17 | 42 | 56.8 | 9 | 4 US-08-197-484-143 | Sequence 143, App |
| 18 | 42 | 56.8 | 9 | 4 US-08-627-820-4 | Sequence 4, Appli |
| 19 | 42 | 56.8 | 9 | 5 PCT-US95-02121-84 | Sequence 84, Appl |
| 20 | 42 | 56.8 | 9 | 5 PCT-US95-02121-143 | Sequence 143, App |
| 21 | 42 | 56.8 | 10 | 1 US-08-465-167A-42 | Sequence 42, Appl |
| 22 | 41 | 55.4 | 9 | 1 US-08-465-167A-8 | Sequence 8, Appli |
| 23 | 41 | 55.4 | 9 | 1 US-08-465-167A-27 | Sequence 27, Appl |
| 24 | 41 | 55.4 | 9 | 4 US-08-197-484-93 | Sequence 93, Appl |
| 25 | 41 | 55.4 | 9 | 4 US-08-197-484-152 | Sequence 152, App |
| 26 | 41 | 55.4 | 9 | 4 US-08-627-820-8 | Sequence 8, Appli |
| 27 | 41 | 55.4 | 9 | 4 US-09-543-608A-44 | Sequence 44, Appl |

| | | | | | |
|----|------|------|------|------------------------|-------------------|
| 28 | 41 | 55.4 | 9 | 5 PCT-US95-02121-93 | Sequence 93, Appl |
| 29 | 41 | 55.4 | 9 | 5 PCT-US95-02121-152 | Sequence 152, App |
| 30 | 41 | 55.4 | 157 | 4 US-09-328-352-5371 | Sequence 5371, Ap |
| 31 | 39 | 52.7 | 9 | 1 US-08-465-167A-33 | Sequence 33, Appl |
| 32 | 39 | 52.7 | 213 | 4 US-09-328-352-7428 | Sequence 7428, Ap |
| 33 | 39 | 52.7 | 447 | 4 US-09-252-991A-23030 | Sequence 23030, A |
| 34 | 38 | 51.4 | 591 | 4 US-09-540-236-3375 | Sequence 3375, Ap |
| 35 | 37.5 | 50.7 | 355 | 4 US-09-345-473E-26 | Sequence 26, Appl |
| 36 | 37 | 50.0 | 9 | 1 US-08-465-167A-26 | Sequence 26, Appl |
| 37 | 37 | 50.0 | 9 | 1 US-08-465-167A-31 | Sequence 31, Appl |
| 38 | 37 | 50.0 | 10 | 1 US-08-465-167A-17 | Sequence 17, Appl |
| 39 | 37 | 50.0 | 10 | 1 US-08-465-167A-46 | Sequence 46, Appl |
| 40 | 37 | 50.0 | 10 | 3 US-08-159-339A-602 | Sequence 602, App |
| 41 | 37 | 50.0 | 10 | 4 US-08-627-820-17 | Sequence 17, Appl |
| 42 | 36 | 48.6 | 275 | 4 US-09-468-433C-24 | Sequence 24, Appl |
| 43 | 36 | 48.6 | 917 | 4 US-09-049-698-41 | Sequence 41, Appl |
| 44 | 36 | 48.6 | 1264 | 1 US-07-789-915A-6 | Sequence 6, Appli |
| 45 | 36 | 48.6 | 1264 | 1 US-08-005-002C-6 | Sequence 6, Appli |

ALIGNMENTS

RESULT 1
US-08-465-167A-1
; Sequence 1, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-1

Query Match 100.0%; Score 74; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 VKVLEYVIKVSARVRF 16
      |||
Db     26 VKVLEYVIKVSARVRF 41

RESULT 2
US-08-627-820-1
; Sequence 1, Application US/08627820
; Patent No. 6464980
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
;             Livingston, Brian D.
;             Sette, Alessandro D.
;             Sidney, John C.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
;                     COMPLETE MAGE 1 GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,820
; FILING DATE: 02-Apr-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,623
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-08-627-820-1
Query Match      100.0%; Score 74; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VKVLEYVIKVSARVRF 16
      |||
Db     26 VKVLEYVIKVSARVRF 41

RESULT 3
US-08-465-167A-24
; Sequence 24, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL

```

```

; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-167A-24
Query Match      100.0%; Score 74; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VKVLEYVIKVSARVRF 16
      |||
Db     277 VKVLEYVIKVSARVRF 292

RESULT 4
US-08-993-118-10
; Sequence 10, Application US/08993118
; Patent No. 5997872
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie;
; APPLICANT: DE SMET, Charles;
; APPLICANT: BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
; TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,118
; FILING DATE:

```



```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
US-08-993-118-10
Query Match 100.0%; Score 74; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
Db 277 VKVLEYVIKVSARVRF 292

RESULT 5
US-08-845-528C-10
; Sequence 10, Application US/08845528C
; Patent No. 6027924
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie;
; APPLICANT: DE SMET, Charles;
; APPLICANT: BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
; TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,528C
; FILING DATE: April 25, 1997
; CLASSIFICATION: 4335
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
US-08-845-528C-10
Query Match 100.0%; Score 74; DB 3; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
```

```

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
Db 277 VKVLEYVIKVSARVRF 292

RESULT 6
US-08-627-820-24
; Sequence 24, Application US/08627820
; Patent No. 6464980
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; Livingston, Brian D.
; Sette, Alessandro D.
; Sidney, John C.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
; COMPLETE MAGE 1 GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,820
; FILING DATE: 02-Apr-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,623
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-08-627-820-24
Query Match 100.0%; Score 74; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
Db 277 VKVLEYVIKVSARVRF 292

RESULT 7
US-09-066-281B-10
; Sequence 10, Application US/09066281B
; Patent No. 6475783
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; DE SMET, Charles; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING
; TITLE OF INVENTION: FOR TUMOR REJECTION ANTIGEN PRECURSOR MAGE-C1 AND MAGE-C2
; TITLE OF INVENTION: AND USES THEREOF
```

```

; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,281B
; FILING DATE: April 24, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455.2 US - JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
US-09-066-281B-10

Query Match 100.0%; Score 74; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
   |||||
Db 277 VKVLEYVIKVSARVRF 292

RESULT 8
US-09-468-433C-10
; Sequence 10, Application US/09468433C
; Patent No. 6680191
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING FOR
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C AN
; TITLE OF INVENTION: MAGE-B FAMILIES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/468,433C
; FILING DATE: December 17, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/066,281
; FILING DATE: April 24, 1998

```

```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-15

Query Match 62.2%; Score 46; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YVIKVSARVR 15
Db 1 YVIKVSARVR 10

RESULT 11
US-08-465-167A-49
; Sequence 49, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

;
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-49

Query Match 62.2%; Score 46; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LEYVIKVSAR 13
Db 1 LEYVIKVSAR 10

RESULT 12
US-08-159-339A-595
; Sequence 595, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-005030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 595:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids

```


/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-159-339A-595

Query Match 62.2%; Score 46; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YVIKVSARVR 15
| | | | |
Db 1 YVIKVSARVR 10

RESULT 13

US-08-627-820-15
; Sequence 15, Application US/08627820
; Patent No. 6464980
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; Livingston, Brian D.
; Sette, Alessandro D.
; Sidney, John C.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
; COMPLETE MAGE 1 GENE

NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: One Market Plaza, Steuart Street Tower
CITY: San Francisco
STATE: CA

COUNTRY: U.S.A.

ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/627,820

FILING DATE: 02-Apr-1996

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/103,623

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14137-60

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 467-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Query Match 62.2%; Score 46; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YVIKVSARVR 15
| | | | |
Db 1 YVIKVSARVR 10

RESULT 14

US-08-465-167A-39
; Sequence 39, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; Livingston, Brian D.
; Sette, Alessandro D.
; Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/465,167A

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/103,623

FILING DATE: 06-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14137-60-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-467-9600

TELEFAX: 415-576-0300

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-465-167A-39

Query Match 60.8%; Score 45; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VLEYVIKVSA 12
| | | | |
Db 1 VLEYVIKVSA 10

RESULT 15

US-08-465-167A-4
; Sequence 4, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; Livingston, Brian D.
; Sette, Alessandro D.
; Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA

Query Match 62.2%; Score 46; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YVIKVSARVR 15
| | | | |
Db 1 YVIKVSARVR 10

```

; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-167A-4

```

```

Query Match      56.8%; Score 42; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 KVLEYVIKV 10
      |||||
Db      1 KVLEYVIKV 9

```

Search completed: October 7, 2004, 15:19:25
Job time : 18.4545 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:17:11 ; Search time 18.5455 Seconds
(without alignments)
82.989 Million cell updates/sec

Title: US-09-336-091-3
Perfect score: 74
Sequence: 1 VKVLEYVIKVSARVRF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 56 | 75.7 | 317 | 2 I38661 | melanoma antigen M |
| 2 | 40 | 54.1 | 242 | 2 D72218 | conserved hypothet |
| 3 | 40 | 54.1 | 542 | 2 E86628 | prophage psl prote |
| 4 | 40 | 54.1 | 594 | 2 F72265 | conserved hypothet |
| 5 | 39 | 52.7 | 379 | 2 T04645 | hypothetical prote |
| 6 | 39 | 52.7 | 444 | 2 H82992 | aminopeptidase P p |
| 7 | 39 | 52.7 | 530 | 2 D70476 | DNA helicase - Aqu |
| 8 | 38.5 | 52.0 | 341 | 2 F90084 | hypothetical prote |
| 9 | 38 | 51.4 | 357 | 2 H71122 | hypothetical prote |
| 10 | 38 | 51.4 | 390 | 2 T10231 | anther-specific pr |
| 11 | 38 | 51.4 | 493 | 2 H70479 | leucine aminopepti |
| 12 | 38 | 51.4 | 761 | 2 D70447 | tetrahydropteroylt |
| 13 | 38 | 51.4 | 846 | 2 H70871 | hypothetical prote |
| 14 | 38 | 51.4 | 869 | 2 A86983 | conserved hypothet |
| 15 | 38 | 51.4 | 869 | 2 S72760 | ppsl protein - Myc |
| 16 | 37.5 | 50.7 | 355 | 2 G96813 | hypothetical prote |
| 17 | 37 | 50.0 | 129 | 2 E90331 | hypothetical prote |
| 18 | 37 | 50.0 | 139 | 2 C70114 | ribosomal protein |
| 19 | 37 | 50.0 | 206 | 2 C96540 | hypothetical prote |
| 20 | 37 | 50.0 | 266 | 2 A97145 | ABC-type iron (III |
| 21 | 37 | 50.0 | 324 | 2 F69758 | transcription regu |
| 22 | 37 | 50.0 | 405 | 2 D90166 | hypothetical prote |
| 23 | 37 | 50.0 | 407 | 2 D86973 | probable chorismat |
| 24 | 37 | 50.0 | 543 | 2 G83825 | hypothetical prote |
| 25 | 37 | 50.0 | 544 | 2 S54531 | hypothetical prote |
| 26 | 37 | 50.0 | 664 | 2 B81269 | hypothetical prote |
| 27 | 37 | 50.0 | 946 | 2 T01460 | hypothetical prote |
| 28 | 37 | 50.0 | 1070 | 1 A54600 | 1-phosphatidylinos |
| 29 | 37 | 50.0 | 1253 | 2 F86436 | hypothetical prote |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 30 | 37 | 50.0 | 1738 | 2 C84507 | hypothetical prote |
| 31 | 37 | 50.0 | 1887 | 2 S61703 | fatty-acid synthas |
| 32 | 37 | 50.0 | 2944 | 2 A54849 | collagen alpha 1(V |
| 33 | 36.5 | 49.3 | 646 | 2 T33346 | hypothetical prote |
| 34 | 36 | 48.6 | 89 | 2 T17491 | hypothetical prote |
| 35 | 36 | 48.6 | 159 | 2 B72764 | hypothetical prote |
| 36 | 36 | 48.6 | 200 | 2 A96664 | unknown protein T1 |
| 37 | 36 | 48.6 | 216 | 2 AF2026 | hypothetical prote |
| 38 | 36 | 48.6 | 314 | 2 JC2360 | melanoma antigen M |
| 39 | 36 | 48.6 | 346 | 2 AH3293 | holliday junction |
| 40 | 36 | 48.6 | 350 | 2 AI3387 | beta-N-acetylhexos |
| 41 | 36 | 48.6 | 355 | 2 T24822 | hypothetical prote |
| 42 | 36 | 48.6 | 378 | 2 F90323 | coenzyme PQQ synth |
| 43 | 36 | 48.6 | 378 | 2 D90347 | coenzyme PQQ synth |
| 44 | 36 | 48.6 | 392 | 2 S72753 | hypothetical prote |
| 45 | 36 | 48.6 | 429 | 2 AB2456 | glycosyltransferas |

ALIGNMENTS

RESULT 1
I38661
melanoma antigen MAGE-4 - human
N;Alternate names: MAGE 41 protein; melanoma antigen MAGE-X2
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1996 #sequence revision 07-Jun-1996 #text change 18-Feb-2000
C;Accession: I38661; I38662; PH1297; PH1298; JC2359; G01446
R;De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; B
oon, T.
Immunogenetics 40, 360-369, 1994
A;Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE fa
A;Reference number: I38659; MUID:95012457; PMID:7927540
A;Accession: I38661
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-317 <DEP1>
A;Cross-references: EMBL:U10687; NID:G5333514; PIDN:AAA68871.1; PID:G5333515
A;Experimental source: antigen MAGE-4a
A;Accession: I38662
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-172,'T',174-317 <DEP2>
A;Cross-references: EMBL:U10688; NID:G5333516; PIDN:AAA68872.1; PID:G5333517
A;Experimental source: antigen MAGE-4b
R;Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van pel
J. Exp. Med. 176, 1453-1457, 1992
A;Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolyti
A;Reference number: PH1294; MUID:93018875; PMID:1402688
A;Accession: PH1297
A;Molecule type: DNA
A;Residues: 169-177 <TRAL>
A;Experimental source: antigen MAGE-4
A;Accession: PH1298
A;Molecule type: DNA
A;Residues: 169-172,'T',174-177 <TRA2>
A;Experimental source: antigen MAGE-41
R;Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
Biochem. Biophys. Res. Commun. 202, 549-555, 1994
A;Title: Cloning and analysis of MAGE-1-related genes.
A;Reference number: JC2358; MUID:94311935; PMID:8037761
A;Accession: JC2359
A;Molecule type: mRNA
A;Residues: 1-172,'T',174-306,'Q',308-317 <DIN>
A;Cross-references: EMBL:U10340; NID:G499123; PIDN:AAA19007.1; PID:G499124
A;Experimental source: melanoma cell line DM150
C;Genetics:
A;Gene: GDB:MAGEA4; MAGE4; MAGE-X2
A;Cross-references: GDB:331119
A;Map position: Xq28-Xq28
A;Introns: #status absent
C;Superfamily: tumor associated protein MAGE
F;169-177/Region: HLA-A1 binding #status predicted

Query Match 75.7%; Score 56; DB 2; Length 317;
Best Local Similarity 73.3%; Pred. No. 0.018;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVR 15
|||||:|:|:|
Db 285 VKVLEHVVRVNARVR 299

RESULT 2
D72218
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: D72218
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D72218
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-242 <ARN>
A;Cross-references: GB:AE001812; GB:AE000512; NID:g4982302; PIDN:AAD36806.1; PID:g498231
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1741
C;Superfamily: conserved hypothetical protein HI0860

Query Match 54.1%; Score 40; DB 2; Length 242;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
||| | | : | | | |
Db 127 VKVTETVKVSAGTVF 142

RESULT 3
E86628
prophage ps1 protein 05, DNA primase [imported] - Lactococcus lactis subsp. lactis (stra
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: E86628
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: E86628
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-542 <STO>
A;Cross-references: GB:AE005176; PID:g12722868; PIDN:AAK04127.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ps105

Query Match 54.1%; Score 40; DB 2; Length 542;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 VLEYVIKVSARVRF 16
|||||:|:| : |
Db 402 VLEYVLKIALSLNF 415

RESULT 4
F72265
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72265
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: F72265
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-594 <ARN>
A;Cross-references: GB:AE001788; GB:AE000512; NID:g4981893; PIDN:AAD36406.1; PID:g498189
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1334

Query Match 54.1%; Score 40; DB 2; Length 594;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 LEYVIKVSARVR 15
|||:| | | |
Db 288 LEYIIKKGRVR 299

RESULT 5
T04645
hypothetical protein F10N7.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Sep-1999
C;Accession: T04645
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15263
A;Accession: T04645
A;Molecule type: DNA
A;Residues: 1-379 <BEV>
A;Cross-references: EMBL:AL021636
A;Experimental source: cultivar Columbia; BAC clone F10N7
C;Genetics:
A;Map position: 4
A;Introns: 48/1; 221/1; 252/3; 287/1; 326/3
A;Note: F10N7.190
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 52.7%; Score 39; DB 2; Length 379;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSA 12
| | | | | : | | | |
Db 14 VSVLEFIISVSA 25

RESULT 6
H82992
aminopeptidase P PA5224 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: H82992
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim;
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H82992
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: GB:AE004935; GB:AE004091; NID:g9951526; PIDN:AAG08609.1; GSPDB:GN001
A;Experimental source: strain PA01

C;Genetics:
A;Gene: pepp; PA5224
C;Superfamily: aminopeptidase P

Query Match 52.7%; Score 39; DB 2; Length 444;
Best Local Similarity 61.5%; Pred. No. 34;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSAR 13
|||:|:|
Db 181 VKVMRYAAEVSAR 193

RESULT 7
D70476
DNA helicase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C;Accession: D70476
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: D70476
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-530 <AQF>
A;Cross-references: GB:AE000770; NID:g2984274; PIDN:AAC07803.1; PID:g2984278; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: helX

Query Match 52.7%; Score 39; DB 2; Length 530;
Best Local Similarity 46.2%; Pred. No. 41;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSAR 13
||:|:|:|
Db 412 VKIVEYLMKIGLR 424

RESULT 8
F90084
hypothetical protein orf341 [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: F90084
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: F90084
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-341 <DOU>
A;Cross-references: GB:AF165818; NID:gl3794434; PIDN:AAK39809.1; GSPDB:GN00150
C;Genetics:
A;Gene: orf341
A;Map position: 1
A;Genome: nucleomorph
C;Keywords: nucleomorph

Query Match 52.0%; Score 38.5; DB 2; Length 341;
Best Local Similarity 44.4%; Pred. No. 32;
Matches 8; Conservative 5; Mismatches 2; Indels 3; Gaps 1;

Qy 1 VKVLEYVI---KVSARVR 15
:|:|:|:|:|:|
Db 162 IKILEYLIIPHKLKKRIR 179

RESULT 9
H71122
hypothetical protein PH0751 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 02-Feb-2001
C;Accession: H71122
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: H71122
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <KAW>
A;Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29842.1; PID:g3257159
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0751
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 51.4%; Score 38; DB 2; Length 357;
Best Local Similarity 52.4%; Pred. No. 42;
Matches 11; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

Qy 1 VKVL-----EYVIKVSARVR 15
|:|:|:|:|:|:|
Db 287 VRVLGRDVKSEYPIEVSGRVR 307

RESULT 10
T10231
another-specific protein homolog T11111.90 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 12-Jun-2003
C;Accession: T10231
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16992
A;Accession: T10231
A;Molecule type: DNA
A;Residues: 1-390 <BEV>
A;Cross-references: EMBL:AL079347; GSPDB:GN00062; ATSP:T11111.90
A;Experimental source: cultivar Columbia; BAC clone F11111
C;Genetics:
A;Gene: ATSP:T11111.90
A;Map position: 4
A;Introns: 60/1; 249/1
C;Superfamily: chalcone/stilbene synthase

Query Match 51.4%; Score 38; DB 2; Length 390;
Best Local Similarity 53.3%; Pred. No. 46;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSARVR 15
|:|:|:|:|:|:|
Db 345 VYVLEYMLEESKKVR 359

RESULT 11
H70479
leucine aminopeptidase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 03-Mar-2003
C;Accession: H70479
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: H70479
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

| | | | |
|--|---|--|--|
| A;Molecule type: DNA A;Residues: 1-493 <AQF> A;Cross-references: GB:AE000772; NID:g2984299; PIDN:AAC07829.1; PID:g2984306; GB:AE00065 A;Experimental source: strain VF5 C;Genetics: A;Gene: pepA C;Superfamily: Cytosol aminopeptidase | | | |
| Query Match 51.4%; Score 38; DB 2; Length 493; Best Local Similarity 63.6%; Pred. No. 58; Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0; | | | |
| QY | 4 LEYVIKVSARV 14 : : 483 LEYIMKVSSNV 493 | | |
| Db | | | |
| RESULT 12 D70447 tetrahydropteroyltriglutamate methyltransferase - Aquifex aeolicus C;Species: Aquifex aeolicus C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 18-Jun-1999 C;Accession: D70447 R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov V. Nature 392, 353-358, 1998 A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus. A;Reference number: A70300; MUID:98196666; PMID:9537320 A;Accession: D70447 A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Residues: 1-761 <AQF> A;Cross-references: GB:AE000752; NID:g2984021; PIDN:AAC07565.1; PID:g2984022; GB:AE00065 A;Experimental source: strain VF5 C;Genetics: A;Gene: metE C;Superfamily: cobalamin-independent methionine synthase | | | |
| Query Match 51.4%; Score 38; DB 2; Length 761; Best Local Similarity 57.1%; Pred. No. 90; Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0; | | | |
| QY | 1 VKVLEYVIKVSARV 14 : : 309 VKLLELLKVSANL 322 | | |
| Db | | | |
| RESULT 13 H70871 hypothetical protein Rv1461 - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999 C;Accession: H70871 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A;Reference number: A70500; MUID:98295987; PMID:9634230 A;Accession: H70871 A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Residues: 1-846 <COL> A;Cross-references: GB:AL021184; GB:AL123456; NID:g3261498; PIDN:CAA15991.1; PID:e123745 A;Experimental source: strain H37Rv C;Genetics: A;Gene: Rv1461 | | | |
| Query Match 51.4%; Score 38; DB 2; Length 846; Best Local Similarity 50.0%; Pred. No. 1e+02; Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0; | | | |

| | | | |
|---|--|--|--|
| QY | 3 VLEYVIKVSARVRF 16 : :: : 626 VVEIIVKPHARVRY 639 | | |
| Db | | | |
| RESULT 14 A86983 conserved hypothetical protein ML0593 [imported] - Mycobacterium leprae C;Species: Mycobacterium leprae C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001 C;Accession: A86983 R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, eam, M.A.; Rutherford, K.M. Nature 409, 1007-1011, 2001 A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S A;Title: Massive gene decay in the leprosy bacillus. A;Reference number: A86909; MUID:21128732; PMID:11234002 A;Accession: A86983 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-869 <STO> A;Cross-references: GB:AL450380; NID:gl3092788; PIDN:CAC30101.1; GSPDB:GN00147 C;Genetics: A;Gene: ML0593 | | | |
| Query Match 51.4%; Score 38; DB 2; Length 869; Best Local Similarity 50.0%; Pred. No. 1e+02; Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0; | | | |
| QY | 3 VLEYVIKVSARVRF 16 : :: : 649 VVEIIVKPHARVRY 662 | | |
| Db | | | |
| RESULT 15 S72760 pps1 protein - Mycobacterium leprae N;Alternate names: B1496_C2_189 protein C;Species: Mycobacterium leprae C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001 C;Accession: S72760; T11013 R;Smith, D.R.; Robison, K. submitted to the EMBL Data Library, November 1993 A;Description: Mycobacterium leprae cosmid B1496. A;Reference number: S72695 A;Accession: S72760 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-869 <SMI> A;Cross-references: EMBL:U00013; NID:g466868; PIDN:AAA17127.1; PID:g466882 R;Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, September 1997 A;Reference number: Z16918 A;Accession: T11013 A;Status: preliminary; translated from GB/EMBL/DDBJ A;Molecule type: DNA A;Residues: 202-481,'A',483-589 <PAR> A;Cross-references: EMBL:Z99125; NID:g2398683; PIDN:CAB16172.1; PID:e343548; PID:g239870 | | | |
| Query Match 51.4%; Score 38; DB 2; Length 869; Best Local Similarity 50.0%; Pred. No. 1e+02; Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0; | | | |
| QY | 3 VLEYVIKVSARVRF 16 : :: : 649 VVEIIVKPHARVRY 662 | | |
| Db | | | |
| Search completed: October 7, 2004, 15:18:15 Job time : 21.5455 secs | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:13:16 ; Search time 10.9091 Seconds
(without alignments)
76.370 Million cell updates/sec

Title: US-09-336-091-3
Perfect score: 74
Sequence: 1 VKVLEYVIKVSARVRF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 74 | 100.0 | 309 | 1 MAG1_HUMAN | P43355 homo sapien |
| 2 | 56 | 75.7 | 317 | 1 MAG4_HUMAN | P43358 homo sapien |
| 3 | 39 | 52.7 | 1004 | 1 RPOC_OENOE | P95405 oenococcus |
| 4 | 38 | 51.4 | 364 | 1 TRU1_THETN | Q8rah7 thermoanaer |
| 5 | 38 | 51.4 | 493 | 1 AMPA_AQUAE | O67868 aquifex aeo |
| 6 | 38 | 51.4 | 761 | 1 METE_AQUAE | O67606 aquifex aeo |
| 7 | 38 | 51.4 | 846 | 1 YE61_MYCTU | O53152 mycobacteri |
| 8 | 38 | 51.4 | 869 | 1 YE61_MYCLE | Q49689 mycobacteri |
| 9 | 37 | 50.0 | 139 | 1 RS6_BORBU | O51142 borrelia bu |
| 10 | 37 | 50.0 | 407 | 1 AROC_MYCLE | Q9ccs6 mycobacteri |
| 11 | 37 | 50.0 | 455 | 1 TRME_FUSNN | Q8rha2 fusbacteri |
| 12 | 37 | 50.0 | 591 | 1 SYD_PSESM | Q87Y31 pseudomonas |
| 13 | 37 | 50.0 | 1070 | 1 P11B_HUMAN | P42338 homo sapien |
| 14 | 37 | 50.0 | 1887 | 1 FAS2_YEAST | P19097 s fatty aci |
| 15 | 37 | 50.0 | 2944 | 1 CA17_HUMAN | Q02388 homo sapien |
| 16 | 36 | 48.6 | 156 | 1 MOAC_METKA | Q8txv5 methanopyru |
| 17 | 36 | 48.6 | 304 | 1 DDL_BACAA | Q81q29 bacillus an |
| 18 | 36 | 48.6 | 314 | 1 MAG6_HUMAN | P43360 homo sapien |
| 19 | 36 | 48.6 | 346 | 1 RUVB_BRUSE | Q8Yiv5 brucella me |
| 20 | 36 | 48.6 | 346 | 1 RUVB_BRUSE | Q8fz02 brucella su |
| 21 | 36 | 48.6 | 392 | 1 YE62_MYCLE | Q49682 mycobacteri |
| 22 | 36 | 48.6 | 458 | 1 TRME_CLOPE | Q8xh30 clostridium |
| 23 | 36 | 48.6 | 602 | 1 SYD_NEIMA | Q9jt23 neisseria m |
| 24 | 36 | 48.6 | 602 | 1 SYD_NEIMB | Q9k0u5 neisseria m |
| 25 | 36 | 48.6 | 1070 | 1 P11B_RAT | Q92110 rattus norv |
| 26 | 36 | 48.6 | 1272 | 1 UBP2_YEAST | Q01476 saccharomyc |
| 27 | 36 | 48.6 | 1764 | 1 YKB4_YEAST | P34241 saccharomyc |
| 28 | 36 | 48.6 | 1828 | 1 DOC2_MOUSE | Q8c3j5 mus musculu |
| 29 | 36 | 48.6 | 1830 | 1 DOC2_HUMAN | Q92608 homo sapien |
| 30 | 36 | 48.6 | 4705 | 1 FAT2_DROME | Q9vw71 drosophila |
| 31 | 35 | 47.3 | 116 | 1 RBFA_UREPA | Q9pqh0 ureaplasma |
| 32 | 35 | 47.3 | 215 | 1 GTH6_ARATH | Q96266 arabidopsis |
| 33 | 35 | 47.3 | 314 | 1 MAG3_HUMAN | P43357 homo sapien |

| | | | | | | | | | |
|------------|--|-----------|------|---------|------------------------|--|--|--|--------------------|
| RESULT 1 | | | | | ALIGNMENTS | | | | |
| MAG1_HUMAN | | | | | STANDARD; PRT; 309 AA. | | | | |
| ID | MAG1_HUMAN | STANDARD; | PRT; | 309 AA. | | | | | |
| AC | P43355; O00346; | | | | | | | | P46962 saccharomyc |
| DT | 01-NOV-1995 (Rel. 32, Created) | | | | | | | | Q09203 caenorhabdi |
| DT | 01-NOV-1995 (Rel. 32, Last sequence update) | | | | | | | | Q07752 trypanosoma |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update) | | | | | | | | Q8r7b9 t arginine |
| DE | Melanoma-associated antigen 1 (MAGE-1 antigen) (Antigen MZ2-E). | | | | | | | | Q58750 methanococc |
| GN | MAGEA1 OR MAGE1 OR MAGE1A. | | | | | | | | Q51422 pseudomonas |
| OS | Homo sapiens (Human). | | | | | | | | Q88nj4 pseudomonas |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | | | | Q06218 saccharomyc |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | | | | | Q8kch0 chlorobium |
| OX | NCBI_TaxID=9606; | | | | | | | | Q01014 herpesvirus |
| RN | [1] | | | | | | | | Q9vv73 drosophila |
| RP | SEQUENCE FROM N.A. | | | | | | | | Q9krb2 vibrio chol |
| RX | MEDLINE=92086861; PubMed=1840703; | | | | | | | | |
| RA | van der Bruggen P., Traversari C., Chomez P., Lurquin C., de Plaen E., | | | | | | | | |
| RA | van den Eynde B., Knuth A., Boon T.; | | | | | | | | |
| RT | "A gene encoding an antigen recognized by cytolytic T lymphocytes on | | | | | | | | |
| RT | a human melanoma."; | | | | | | | | |
| RL | Science 254:1643-1647(1991). | | | | | | | | |
| RN | [2] | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RC | TISSUE=Skin; | | | | | | | | |
| RX | MEDLINE=94311935; PubMed=8037761; | | | | | | | | |
| RA | Ding M., Beck R.J., Keller C.J., Fenton R.G.; | | | | | | | | |
| RT | "Cloning and analysis of MAGE-1-related genes."; | | | | | | | | |
| RL | Biochem. Biophys. Res. Commun. 202:549-555(1994). | | | | | | | | |
| RN | [3] | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RX | MEDLINE=20314869; PubMed=10854409; | | | | | | | | |
| RA | Mallon A.M., Platzer M., Bates R., Gloeckner G., Botcherby M., | | | | | | | | |
| RA | Nordsiek G., Strivens M.A., Kioschis P., Dangel A., Cunningham D., | | | | | | | | |
| RA | Straw R., Weston P., Hunter C., Gilbert M., Fernando S., Goodall K., | | | | | | | | |
| RA | Kerry G., Greystrom J.S., Clark D., Goerdes M., Blechschmidt K., | | | | | | | | |
| RA | Rump A., Hinzmann B., Mundy C.R., Miller W., Poustka A., Herman G.E., | | | | | | | | |
| RA | Rhodes M., Denny P., Rosenthal A., Brown S.D.M.; | | | | | | | | |
| RT | "Comparative genome sequence analysis of the Bpa/Str region in mouse | | | | | | | | |
| RT | and man."; | | | | | | | | |
| RL | Genome Res. 10:758-775(2000). | | | | | | | | |
| RN | [4] | | | | | | | | |
| RP | SEQUENCE FROM N.A., AND VARIANT ALA-32. | | | | | | | | |
| RA | Chen H., Wang L., Mei M., Qin L., Cong X., Xu J., Wei L., Wang Y., | | | | | | | | |
| RA | Chen W.; | | | | | | | | |
| RT | "The polymorphism of MAGE-1 gene in Chinese people."; | | | | | | | | |
| RL | Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases. | | | | | | | | |
| RN | [5] | | | | | | | | |
| RP | MUTAGENESIS. | | | | | | | | |
| RC | TISSUE=Blood; | | | | | | | | |
| RX | MEDLINE=94157413; PubMed=8113684; | | | | | | | | |
| RA | Gaugler B., van den Eynde B., van der Bruggen P., Romero P., | | | | | | | | |
| RA | Gaforio J.J., de Plaen E., Lethe B., Brasseur F., Boon T.; | | | | | | | | |
| RT | "Human gene MAGE-3 codes for an antigen recognized on a melanoma by | | | | | | | | |
| RT | autologous cytolytic T lymphocytes."; | | | | | | | | |
| RL | J. Exp. Med. 179:921-930(1994). | | | | | | | | |
| RN | [6] | | | | | | | | |

RP SUBCELLULAR LOCATION.
RX MEDLINE=95012905; PubMed=7927954;
RA Schultz-Thater E., Juretic A., Dellabona P., Luscher U., Siegrist W.,
RA Harder F., Heberer M., Zuber M., Spagnoli G.C.;
RT "MAGE-1 gene product is a cytoplasmic protein.";
RL Int. J. Cancer 59:435-439(1994).
CC -!- FUNCTION: Not known, though may play a role in embryonal
CC development and tumor transformation or aspects of tumor
CC progression. Antigen recognized on a melanoma by autologous
CC cytolytic T lymphocytes.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Expressed in many tumors of several types,
CC such as melanoma, head and neck squamous cell carcinoma, lung
CC carcinoma and breast carcinoma, but not in normal tissues except
CC for testes. Never expressed in kidney tumors, leukemias and
CC lymphomas.
CC -!- SIMILARITY: Contains 1 MAGE domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M77481; AAA03229.1; --
DR EMBL; U82670; -; NOT ANNOTATED_CDS.
DR EMBL; AY148486; AAN62752.1; --
DR Genew; HGNC:6796; MAGEA1.
DR MIM; 300016; --
DR GO; GO:0005886; C:plasma membrane; TAS.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Polymorphism; Tumor antigen.
FT DOMAIN 102 301 MAGE.
FT DOMAIN 33 36 POLY-SER.
FT VARIANT 32 32 T -> A (probable polymorphism).
FT VARIANT 72 72 /FTId=VAR_004283.
FT MUTAGEN 163 163 D->A: ABOLISHES HLA-A1 BINDING.
FT MUTAGEN 169 169 Y->A: ABOLISHES HLA-A1 BINDING.
SQ SEQUENCE 309 AA; 34342 MW; 544EEB1F9F4E9D33 CRC64;

Query Match 100.0%; Score 74; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVRF 16
DQ 277 VKVLEYVIKVSARVRF 292
|||||
RESULT 2
MAG4_HUMAN
ID MAG4_HUMAN STANDARD; PRT; 317 AA.
AC P43358;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Melanoma-associated antigen 4 (MAGE-4 antigen) (MAGE-X2) (MAGE-41).
GN MAGEA4 OR MAGE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95012457; PubMed=7927540;
RA de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,

RA de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL Immunogenetics 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=94311935; PubMed=8037761;
RA Ding M., Beck R.J., Keller C.J., Fenton R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL Biochem. Biophys. Res. Commun. 202:549-555(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95369706; PubMed=7642112;
RA Imai Y., Shichijo S., Yamada A., Katayama T., Yano H., Itoh K.;
RT "Sequence analysis of the MAGE gene family encoding human tumor-
RT rejection antigens.";
RL Gene 160:287-290(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Duodenum;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman A., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: Contains 1 MAGE domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10687; AAA68871.1; --
DR EMBL; U10688; AAA68872.1; --
DR EMBL; U10340; AAA19007.1; --
DR EMBL; D32077; BAA06843.1; --
DR EMBL; BC017723; AAH17723.1; --
DR PIR; I38661; I38661.
DR PDB; 1I4F; 25-JUL-01.
DR Genew; HGNC:6802; MAGEA4.
DR MIM; 300175; --
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Polymorphism; Tumor antigen; 3D-structure.
KW


```
FT DOMAIN 110 309 MAGE.
FT DOMAIN 41 44 POLY-SER.
FT VARIANT 173 173 T -> A.
FT /FTId=VAR_004284.
FT CONFLICT 307 307 E -> Q (IN REF. 2).
SQ SEQUENCE 317 AA; 34929 MW; 3CFAC0E2B696257C CRC64;

Query Match 75.7%; Score 56; DB 1; Length 317;
Best Local Similarity 73.3%; Pred. No. 0.0089;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVR 15
Db 285 VKVLEHVVRNARVR 299

RESULT 3
RPOC_OENOE
ID_RPOC_OENOE STANDARD; PRT; 1004 AA.
AC P95405;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
DE beta' chain) (RNA polymerase beta' subunit) (fragment).
GN RPOC.
OS Oenococcus oeni (Leuconostoc oenos).
OC Bacteria; Firmicutes; Lactobacillales; Oenococcus.
OX NCBI_TaxID=1247;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCDO 1674;
RX MEDLINE=97016803; PubMed=8863429;
RA Morse R., Collins M.D., O'Hanlon K., Wallbanks S., Richardson P.T.;
RT "Analysis of the beta' subunit of DNA-dependent RNA polymerase does
RT not support the hypothesis inferred from 16S rRNA analysis that
RT Oenococcus oeni (formerly Leuconostoc oenos) is a tachytelic
RT (fast-evolving) bacterium.";
RL Int. J. Syst. Bacteriol. 46:1004-1009(1996).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: The enzyme consists of the sigma chain and the core
CC enzyme which is composed of 2 alpha chains, 1 beta chain, and 1
CC beta' chain.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X96384; CAA65248.1; -.
DR HSSP; Q9KWU6; 1HQM.
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_pol_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR SMART; SM00663; RPOLA_N; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription.
FT NON_TER 1 1
```

```
FT NON_TER 1004 1004
SQ SEQUENCE 1004 AA; 111965 MW; 73750DF47F3A2C36 CRC64;

Query Match 52.7%; Score 39; DB 1; Length 1004;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYVIKVSARVR 15
Db 942 EYVLPISARLR 952

RESULT 4
TRU1_THETN
ID_TRU1_THETN STANDARD; PRT; 364 AA.
AC Q8RAH7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
DE 1 (EC 2.1.1.61).
GN TRMU1 OR TTE1243.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC homocysteine + tRNA containing 5-methylaminomethyl-2-
CC thiouridylate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the trmU family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE013086; AAM24468.1; -.
DR HAVAP; MF_00144; -.
DR InterPro; IPR004506; TrmU.
DR Pfam; PF03054; tRNA_Me_trans; 1.
DR TIGRFAMs; TIGR00420; trmU; 1.
KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ SEQUENCE 364 AA; 41666 MW; 2DCAC85A65BC071 CRC64;

Query Match 51.4%; Score 38; DB 1; Length 364;
Best Local Similarity 43.8%; Pred. No. 21;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVR 16
Db 296 IEKLEEEIKVTAKIRY 311

RESULT 5
AMPA_AQUAE
ID_AMPA_AQUAE STANDARD; PRT; 493 AA.
AC O67868;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE (LAP) (Leucyl aminopeptidase).
GN PEPA OR AQ_2099.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: Presumably involved in the processing and regular
CC turnover of intracellular proteins. Catalyzes the removal of
CC unsubstituted N-terminal amino acids from various peptides (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
CC Xbb-, in which Xaa is preferably Leu, but may be other amino acids
CC including Pro although not Arg or Lys, and Xbb may be Pro.
CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M17.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000772; AAC07829.1; --
DR PIR; H70479; H70479.
DR HSSP; P00727; 1LAP.
DR MEROPS; M17.UPW; --
DR HAMAP; MF_00181; -; 1.
DR InterPro; IPR000819; Peptidase_M17_C.
DR InterPro; IPR008283; Peptidase_M17_N.
DR Pfam; PF00883; Peptidase_M17; 1.
DR Pfam; PF02789; Peptidase_M17_N; 1.
DR PRINTS; PR00481; LAMNOPPTDASE.
DR PROSITE; PS00631; CYTOSOL_AP; 1.
KW Hydrolase; Aminopeptidase; Manganese; Complete proteome.
FT METAL 257 257 MANGANESE 2 (BY SIMILARITY).
FT METAL 262 262 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 280 280 MANGANESE 2 (BY SIMILARITY).
FT METAL 339 339 MANGANESE 1 (BY SIMILARITY).
FT METAL 341 341 MANGANESE 1 AND 2 (BY SIMILARITY).
FT ACT_SITE 269 269 POTENTIAL.
FT ACT_SITE 343 343 POTENTIAL.
SQ SEQUENCE 493 AA; 54543 MW; A32B499C7A52065B CRC64;

Query Match 51.4%; Score 38; DB 1; Length 493;
Best Local Similarity 63.6%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 LEYVIKVSARV 14
|||:|||||
Db 483 LEYIMKVSSNV 493

RESULT 6
METE_AQUAE STANDARD; PRT; 761 AA.
AC O67606;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase

DE (EC 2.1.1.14) (Methionine synthase, vitamin-B12 independent isozyme)
DE (Cobalamin-independent methionine synthase).
GN METE OR AQ_1710.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: Catalyzes the transfer of a methyl group from 5-
CC methyltetrahydrofolate to homocysteine resulting in methionine
CC formation (By similarity).
CC -!- CATALYTIC ACTIVITY: 5-methyltetrahydropteroyltri-L-glutamate + L-
CC homocysteine = tetrahydropteroyltri-L-glutamate + L-methionine.
CC -!- COFACTOR: Zinc; binds one ion per subunit (By similarity).
CC -!- PATHWAY: Terminal step in the de novo biosynthesis of methionine.
CC -!- SIMILARITY: Belongs to the vitamin-B12 independent methionine
CC synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000752; AAC07565.1; --
DR PIR; D70447; D70447.
DR HAMAP; MF_00172; -; 1.
DR InterPro; IPR006276; Met_syn_B12ind.
DR InterPro; IPR002629; Methionine_synth.
DR Pfam; PF01717; Methionine_synt; 1.
DR ProDom; PD004692; Methionine_synt; 2.
DR TIGRFAMs; TIGR01371; met_syn_B12ind; 1.
KW Transferase; Methyltransferase; Methionine biosynthesis; Zinc; Repeat;
KW Complete proteome.
FT METAL 643 643 ZINC (BY SIMILARITY).
FT METAL 645 645 ZINC (BY SIMILARITY).
FT METAL 729 729 ZINC (BY SIMILARITY).
SQ SEQUENCE 761 AA; 88631 MW; C8F69759F86D5383 CRC64;

Query Match 51.4%; Score 38; DB 1; Length 761;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 VKVLEVYIKVSARV 14
|||:|||||
Db 309 VKLLELLKVSANL 322

RESULT 7
YE61_MYCTU STANDARD; PRT; 846 AA.
ID_YE61_MYCTU
AC O53152;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Rv1461/MT1508/Mb1496 [Contains: Endonuclease PI-
DE MtuHIIP (EC 3.1.-.-) (Mtu pps1 intein)].
GN Rv1461 OR MT1508 OR MTV007.08 OR Mb1496.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;

RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Barrell B.G.;
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
RN [4]
RP This protein undergoes a protein self splicing that involves
RC a post-translational excision of the intervening region (intein)
RX followed by peptide ligation (Potential).
RA -!- SIMILARITY: Belongs to the UPF0051 (ycf24) family.
RN [5]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RX the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RA or send an email to license@isb-sib.ch).
RN [6]
RP EMBL; AL021184; CAA15991.1; -.
DR EMBL; AE007020; AAK45772.1; -.
DR EMBL; BX248339; CAD96163.1; -.
DR PIR; H70871; H70871.
DR HSSP; P17255; 1VDE.
DR REBASE; 4231; PI-MTUHIIP.
DR TIGR; MT1508; -.
DR TubercuList; RV1461; -.
DR InterPro; IPR003587; Hedgehog_hint_N.
DR InterPro; IPR003586; Hedgehog_hint_C.
DR InterPro; IPR007868; Hom_end_hint.
DR InterPro; IPR006142; INTEIN.
DR InterPro; IPR004042; Intein_endonuc.
DR InterPro; IPR006141; Intein_S.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF05203; Hom_end_hint; 1.
DR Pfam; PF01458; UPF0051; 2.
DR PRINTS; PR00379; INTEIN.
DR SMART; SM00305; HintC; 1.
DR SMART; SM00306; HintN; 1.
DR TIGRFAMs; TIGR01443; intein_Cterm; 1.

DR PROSITE; PS50818; INTEIN_C_TER; 1.
DR PROSITE; PS50819; INTEIN_ENDONUCLEASE; 1.
DR PROSITE; PS50817; INTEIN_N_TER; 1.
KW Hypothetical protein; Protein splicing; Autocatalytic cleavage;
KW Hydrolase; Nuclease; Endonuclease; Intron homing; Complete proteome.
FT CHAIN 1 252 HYPOTHETICAL PROTEIN RV1461, 1ST PART
(POTENTIAL).
FT CHAIN 253 611 ENDONUCLEASE PI-MTUHIIP (POTENTIAL).
FT CHAIN 612 846 HYPOTHETICAL PROTEIN RV1461, 2ND PART
(POTENTIAL).
SQ SEQUENCE 846 AA; 94171 MW; 468CEEF979B02222 CRC64;
Query Match 51.4%; Score 38; DB 1; Length 846;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 3 VLEYVIKVSARVRF 16
Db 626 VVEIIVKPHARVRY 639
RESULT 8
YE61_MYCLE STANDARD; PRT; 869 AA.
ID YE61_MYCLE O33141;
AC Q49689; O33141;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ML0593 [Contains: Mle pps1 intein].
GN ML0593 OR MLCL536.28C OR B1496_C2_189.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -!- PTM: This protein undergoes a protein self splicing that involves
CC a post-translational excision of the intervening region (intein)
CC followed by peptide ligation (Potential).
CC -!- SIMILARITY: Belongs to the UPF0051 (ycf24) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U00013; AAA17127.1; -.
DR EMBL; Z99125; CAB16171.1; -.
DR EMBL; Z99125; CAB16172.1; -.
DR EMBL; AL583919; CAC30101.1; -.
DR PIR; A86983; A86983.
DR PIR; S72760; S72760.
DR Leprona; ML0593; -.

DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
KW Lyase; Aromatic amino acid biosynthesis; Complete proteome.
SQ SEQUENCE 407 AA; 42473 MW; 1F677D07ADB556AF CRC64;

Query Match 50.0%; Score 37; DB 1; Length 407;
Best Local Similarity 43.8%; Pred. No. 37;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VKVLEYVIVKVSARVRF 16
Db 161 VEVLSHVISIGASARY 176

RESULT 11

TRME_FUSNN
ID TRME_FUSNN STANDARD; PRT; 455 AA.
AC Q8RHA2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE tRNA modification GTPase trmE.
GN TRME OR FN0006.

OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586.";

RL J. Bacteriol. 184:2005-2018(2002).
CC -!- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate.
CC Involved in the biosynthesis of the hypermodified nucleoside 5-
CC methylaminomethyl-2-thiouridine, which is found in the wobble
CC position of some tRNAs (By similarity).
CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC proteins. TrmE subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE010516; AAL94219.1; -.
DR HAMAP; MF 00379; -; 1.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR002917; MMR_HSR1.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR004520; ThdF.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00326; GTP1_OBG.
DR TIGRFAMS; TIGR00650; MG442; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR TIGRFAMS; TIGR00450; thdF; 1.
KW tRNA processing; GTP-binding; Complete proteome.
FT NP_BIND 226 233 GTP (POTENTIAL).
FT NP_BIND 273 277 GTP (POTENTIAL).
FT NP_BIND 333 336 GTP (POTENTIAL).
SQ SEQUENCE 455 AA; 50630 MW; 8BE4CE1F983B76B3 CRC64;

Query Match 50.0%; Score 37; DB 1; Length 455;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KVLEYVIVKVSARV 14
Db 96 KVLEVVLKNGARI 108

RESULT 12

SYD_PSESM
ID SYD_PSESM STANDARD; PRT; 591 AA.
AC Q87Y31;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (Asprs).
GN ASPs OR PSPTO3981.

OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DC3000;

RX MEDLINE=22834015; PubMed=12928499;

RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidse T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,
RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
Pseudomonas syringae pv. tomato DC3000.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).

CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +

CC diphosphate + L-aspartyl-tRNA(Asp).

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE016870; AAO57440.1; -.

DR TIGR; PSPTO3981; -.

DR HAMAP; MF 00044; -; 1.

DR InterPro; IPR004115; GAD_dom.

DR InterPro; IPR004364; tRNA-synt_2.

DR InterPro; IPR002312; tRNA-synt_asp.

DR InterPro; IPR004365; tRNA_anti.

DR Pfam; PF02938; GAD; 1.

DR Pfam; PF00152; tRNA-synt_2; 1.

DR Pfam; PF01336; tRNA_anti; 1.

DR PRINTS; PR01042; TRNASYNTHASP.

DR PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.

KW Aminoacyl-tRNA synthetase; protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.

SQ SEQUENCE 591 AA; 66351 MW; 4A841555305F8F36 CRC64;

Query Match

Best Local Similarity 50.0%; Score 37; DB 1; Length 591;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL Nature 387:103-105(1997).
RN [4]
RP MUTAGENESIS OF GLY-1250.
RC STRAIN=S288c;
RX MEDLINE=94316198; PubMed=8041367;
RA Inokoshi J., Tomoda H., Hashimoto H., Watanabe A., Takeshima H.,
RA Omura S.;
RT "Cerulenin-resistant mutants of Saccharomyces cerevisiae with an
altered fatty acid synthase gene.";
RL Mol. Gen. Genet. 244:90-96(1994).
CC -!- FUNCTION: Fatty acid synthetase catalyzes the formation of
long-chain fatty acids from acetyl-CoA, malonyl-CoA and NADPH.
CC The alpha subunit contains domains for: acyl carrier protein,
CC 3-oxoacyl-[acyl-carrier protein] reductase, and 3-oxoacyl-[acyl-
CC carrier-protein] synthase. This subunit coordinates the binding
CC of the six beta subunits to the enzyme complex.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
CC long-chain acyl-CoA + N CoA + N CO(2) + 2N NADP(+).
CC -!- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
CC [acyl-carrier protein].
CC -!- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -!- SUBUNIT: [Alpha(6)beta(6)] hexamers of two multifunctional
CC subunits (alpha and beta).
CC -!- SIMILARITY: TO THE FATTY ACID SYNTHETASE, SUBUNIT ALPHA FROM
CC OTHER FUNGI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03936; AAA34601.1; -;
DR EMBL; X76890; CAA54218.1; -;
DR EMBL; X94561; CAA64256.1; -;
DR EMBL; Z73586; CAA97947.1; -;
DR EMBL; Z73587; CAA97948.1; -;
DR PIR; S61703; S61703.
DR GerOnline; 144213; -;
DR SGD; S0006152; FAS2.
DR InterPro; IPR008278; 4-PPT transf.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR004568; Pantethn_trn.
DR InterPro; IPR006162; Ppantne_S.
DR Pfam; PF01648; ACPS; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt C; 1.
DR TIGRFAMs; TIGR00556; pantethn_trn; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 1.
KW Fatty acid biosynthesis; Multifunctional enzyme; Oxidoreductase;
KW Transferrase; NADP; Phosphopantetheine.
FT DOMAIN 1 ? ACYL CARRIER (ACP).
FT DOMAIN 675 874 BETA-KETOACYL REDUCTASE.
FT DOMAIN 1149 1363 BETA-KETOACYL SYNTHASE.
FT BINDING 180 180 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT ACT_SITE 1305 1305 BETA-KETOACYL SYNTHASE (BY SIMILARITY).

FT MUTAGEN 1250 1250 G->S: CERULENIN-RESISTANCE.
FT CONFLICT 310 310 G -> GTTGTGG (IN REF. 1).
FT CONFLICT 594 594 T -> I (IN REF. 1).
FT CONFLICT 941 1019 AKLRKELVETSEVRKAVSIETALEHKVVGNSADAAYAQVE
FT IQPRANIQLDPFELKPKYQVKQIAPAELEGLLDLERVI ->
FT CLNCVKSWLKLKLERQFPKLLWSIRLSMAIALMLHMLKS
FT KFNQELTFNWTSQNRNHTNRLNKLPLSLRVCWIKELF
FT (IN REF. 1).
FT CONFLICT 1036 1041 RWMEA -> KMGNGS (IN REF. 1).
FT CONFLICT 1408 1408 A -> S (IN REF. 1).
FT CONFLICT 1671 1671 N -> T (IN REF. 1).
SQ SEQUENCE 1887 AA; 206946 MW; 08B872734CF3AEEA CRC64;
Query Match 50.0%; Score 37; DB 1; Length 1887;
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 EYVIKVSAR 13
Db 1672 EYVAKVSAR 1680
RESULT 15
CA17_HUMAN
ID CA17_HUMAN STANDARD; PRT; 2944 AA.
AC Q02388; Q14054; Q16507;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Collagen alpha 1(VII) chain precursor (long-chain collagen) (LC
DE collagen).
GN COL7A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94327588; PubMed=8051117;
RA Christiano A.M., Greenspan D.S., Lee S., Vitto J.;
RT "Cloning of human type VII collagen. Complete primary sequence of the
RT alpha 1(VII) chain and identification of intragenic polymorphisms.";
RL J. Biol. Chem. 269:20256-20262(1994).
RN [2]
RP SEQUENCE OF 128-1493 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93338437; PubMed=1307247;
RA Christiano A.M., Rosenbaum L.M., Chung-Honet L.C., Parente M.G.,
RA Woodley D.T., Pan T.C., Zhang R.Z., Chu M.-L., Burgeson R.E.,
RA Vitto J.;
RT "The large non-collagenous domain (NC-1) of type VII collagen is
RT amino-terminal and chimeric. Homology to cartilage matrix protein,
RT the type III domains of fibronectin and the A domains of von
RT Willebrand factor.";
RL Hum. Mol. Genet. 1:475-481(1992).
RN [3]
RP SEQUENCE OF 815-1439 FROM N.A.
RX MEDLINE=91334380; PubMed=1871109;
RA Parente M.G., Chung L.C., Rynaenen J., Woodley D.T., Wynn K.W.,
RA Bauer E.A., Mattei M.-G., Chu M.-L., Vitto J.;
RT "Human type VII collagen: cDNA cloning and chromosomal mapping of the
RT gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:6931-6935(1991).
RN [4]
RP SEQUENCE OF 369-1255 FROM N.A.
RX MEDLINE=93107742; PubMed=1469284;
RA Gammon W.R., Abernethy M.L., Padilla K.M., Prisayan P.S.,
RA Cook M.E., Wright J., Briggaman R.A., Hunt S.W. III;
RT "Noncollagenous (NC1) domain of collagen VII resembles multidomain
RT adhesion proteins involved in tissue-specific organization of
RT extracellular matrix.";
RL J. Invest. Dermatol. 99:691-696(1992).
RN [5]
RP SEQUENCE OF 340-675 FROM N.A.

RC TISSUE=Keratinocytes;
RX MEDLINE=92231902; PubMed=1567409;
RA Tanaka T., Takahashi K., Furukawa F., Imamura S.;
RT "Molecular cloning and characterization of type VII collagen cDNA.";
RL Biochem. Biophys. Res. Commun. 183:958-963(1992).
RN [6]
RP SEQUENCE OF 2395-2944 FROM N.A.
RX MEDLINE=93271985; PubMed=8499916;
RA Greenspan D.S.;
RT "The carboxyl-terminal half of type VII collagen, including the non-
RT collagenous NC-2 domain and intron/exon organization of the
RT corresponding region of the COL7A1 gene.";
RL Hum. Mol. Genet. 2:273-278(1993).
RN [7]
RP SEQUENCE OF 1-87 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94375010; PubMed=8088784;
RA Christiano A.M., Hoffman G.G., Chung-Honet L.C., Lee S., Cheng W.,
RA Uitto J., Greenspan D.S.;
RT "Structural organization of the human type VII collagen gene (COL7A1),
RT composed of more exons than any previously characterized gene.";
RL Genomics 21:169-179(1994).
RN [8]
RP VARIANT EBDSC ARG-2034.
RX MEDLINE=89227237; PubMed=26533224;
RA Fine J.D., Johnson L., Wright T.;
RT "Epidermolysis bullosa simplex superficialis. A new variant of
RT epidermolysis bullosa characterized by subcorneal skin cleavage
RT mimicking peeling skin syndrome.";
RL Arch. Dermatol. 125:633-638(1989).
RN [9]
RP REVIEW ON DEB VARIANTS.
RX MEDLINE=98041696; PubMed=9375848;
RA Jaervikallio A., Pulkkinen L., Uitto J.;
RT "Molecular basis of dystrophic epidermolysis bullosa: mutations in
RT the type VII collagen gene (COL7A1).";
RL Hum. Mutat. 10:338-347(1997).
RN [10]
RP VARIANT DEB LYS-2798.
RX MEDLINE=93291877; PubMed=8513326;
RA Christiano A.M., Greenspan D.S., Hoffman G.G., Zhang X., Tamai Y.,
RA Lin A.N., Dietz H.C., Hovnanian A., Uitto J.;
RT "A missense mutation in type VII collagen in two affected siblings
RT with recessive dystrophic epidermolysis bullosa.";
RL Nat. Genet. 4:62-66(1993).
RN [11]
RP VARIANT DEB SER-2040.
RX MEDLINE=94224777; PubMed=8170945;
RA Christiano A.M., Ryynaenen M., Uitto J.;
RT "Dominant dystrophic epidermolysis bullosa: identification of a
RT Gly-->Ser substitution in the triple-helical domain of type VII
RT collagen.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:3549-3553(1994).
RN [12]
RP VARIANT DEB CYS-2623.
RX MEDLINE=96081220; PubMed=8541842;
RA Christiano A.M., Lee J.Y.-Y., Chen W.J., Laforgia S., Uitto J.;
RT "Pretibial epidermolysis bullosa: genetic linkage to COL7A1 and
RT identification of a glycine-to-cysteine substitution in the triple-
RT helical domain of type VII collagen.";
RL Hum. Mol. Genet. 4:1579-1583(1995).
RN [13]
RP VARIANT DEB ARG-2043.
RX MEDLINE=95164985; PubMed=7861014;
RA Christiano A.M., Morricone A., Paradisi M., Angelo C., Mazzanti C.,
RA Cavalieri R., Uitto J.;
RT "A glycine-to-arginine substitution in the triple-helical domain of
RT type VII collagen in a family with dominant dystrophic epidermolysis
RT bullosa.";
RL J. Invest. Dermatol. 104:438-440(1995).
RN [14]
RP VARIANTS DEB.
RX MEDLINE=96220218; PubMed=8644729;

RA Christiano A.M., McGrath J.A., Tan K.C., Uitto J.;
RT "Glycine substitutions in the triple-helical region of type VII
RT collagen result in a spectrum of dystrophic epidermolysis bullosa
RT phenotypes and patterns of inheritance.";
RL Am. J. Hum. Genet. 58:671-681(1996).
RN [15]
RP VARIANT DEB ARG-2575.
RX MEDLINE=96154068; PubMed=8592061;
RA Shimizu H., McGrath J.A., Christiano A.M., Nishikawa T., Uitto J.;
RT "Molecular basis of recessive dystrophic epidermolysis bullosa:
RT genotype/phenotype correlation in a case of moderate clinical
RT severity.";
RL J. Invest. Dermatol. 106:119-124(1996).
RN [16]
RP VARIANT DEB ARG-1782.
RX MEDLINE=96183562; PubMed=8618018;
RA Christiano A.M., McGrath J.A., Uitto J.;
RT "Influence of the second COL7A1 mutation in determining the
RT phenotypic severity of recessive dystrophic epidermolysis bullosa.";
RL J. Invest. Dermatol. 106:766-770(1996).
RN [17]
RP VARIANT DEB ASP-2073.
RX MEDLINE=96310789; PubMed=8757758;
RA Dunnill M.G.S., McGrath J.A., Richards A.J., Christiano A.M.,
RA Uitto J., Pope F.M., Eady R.A.J.;
RT "Clinicopathological correlations of compound heterozygous COL7A1
RT mutations in recessive dystrophic epidermolysis bullosa.";
RL J. Invest. Dermatol. 107:171-177(1996).
RN [18]
RP VARIANTS DEB TRP-1982; GLY-2008; ALA-2025; GLU-2049; TRP-2063 AND
RP ARG-2575.
RX MEDLINE=97465605; PubMed=9326325;
RA Hovnanian A., Rochat A., Bodemer C., Petit E., Rivers C.A., Prost C.,
RA Fraitag S., Christiano A.M., Uitto J., Lathrop M., Barrandon Y.,
RA de Prost Y.;
RT "Characterization of 18 new mutations in COL7A1 in recessive
RT dystrophic epidermolysis bullosa provides evidence for distinct
RT molecular mechanisms underlying defective anchoring fibril
RT formation.";
RL Am. J. Hum. Genet. 61:599-610(1997).
RN [19]
RP VARIANT DEB ARG-1652.
RX MEDLINE=98106792; PubMed=9444387;
RA Cserhalmi-Friedman P.B., Karpati S., Horvath A., Christiano A.M.;
RT "Identification of a glycine substitution and a splice site mutation
RT in the type VII collagen gene in a proband with mitis recessive
RT dystrophic epidermolysis bullosa.";
RL Arch. Dermatol. Res. 289:640-645(1997).
RN [20]
RP VARIANTS DEB ARG-2009 AND ARG-2043.
RX MEDLINE=97358588; PubMed=9215684;
RA Winberg J.-O., Hammami-Hauasli N., Nilssen O., Anton-Lamprecht I.,
RA Naylor S.L., Kerbacher K., Zimmermann M., Krajci P.,
RA Gedde-Dahl T. Jr., Bruckner-Tuderman L.;
RT "Modulation of disease severity of dystrophic epidermolysis bullosa by
RT a splice site mutation in combination with a missense mutation in the
RT COL7A1 gene.";
RL Hum. Mol. Genet. 6:1125-1135(1997).
RN [21]
RP VARIANTS DEB ASP-1519; ASP-2006; GLU-2015 AND ARG-2034.
RX MEDLINE=98334662; PubMed=9668111;
RA Hammami-Hauasli N., Schumann H., Raghunath M., Kilgus O., Luethi U.,
RA Luger T., Bruckner-Tuderman L.;
RT "Some, but not all, glycine substitution mutations in COL7A1 result in
RT intracellular accumulation of collagen VII, loss of anchoring
RT fibrils, and skin blistering.";
RL J. Biol. Chem. 273:19228-19234(1998).
RN [22]
RP VARIANTS DEB CYS-2008; ARG-2207 AND SER-2775.
RX MEDLINE=98410969; PubMed=9740253;
RA Kon A., Pulkkinen L., Ishida-Yamamoto A., Hashimoto I., Uitto J.;
RT "Novel COL7A1 mutations in dystrophic forms of epidermolysis
RT bullosa.";

RL J. Invest. Dermatol. 111:534-537(1998).
RN [23]
RP VARIANT DEB ARG-1347.
RX MEDLINE=99019477; PubMed=9804332;
RA Terracina M., Posteraro P., Schubert M., Sonogo G., Atzori F.,
Query Match 50.0%; Score 37; DB 1; Length 2944;
Best Local Similarity 63.6%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 4 LEYVIKVSARV 14
Db 569 LSYTVRVSARV 579

Search completed: October 7, 2004, 15:13:59
Job time : 13.9091 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:14:05 ; Search time 57.8182 Seconds
(without alignments)
87.313 Million cell updates/sec

Title: US-09-336-091-3 ~
Perfect score: 74
Sequence: 1 VKVLEYVIKVSARVRF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 74 | 100.0 | 316 | 4 Q8WWH6 | Q8wwh6 homo sapien |
| 2 | 56 | 75.7 | 317 | 4 Q14798 | Q14798 homo sapien |
| 3 | 56 | 75.7 | 318 | 4 Q9BUN9 | Q9bun9 homo sapien |
| 4 | 42 | 56.8 | 248 | 16 Q8KEI5 | Q8kei5 chlorobium |
| 5 | 41 | 55.4 | 152 | 17 Q8ZZ01 | Q8zz01 pyrobaculum |
| 6 | 41 | 55.4 | 1390 | 16 Q8EA16 | Q8eal6 shewanella |
| 7 | 40 | 54.1 | 242 | 16 Q9X263 | Q9x263 thermotoga |
| 8 | 40 | 54.1 | 274 | 5 Q8WRT9 | Q8wrt9 trypanosoma |
| 9 | 40 | 54.1 | 542 | 9 Q9AZI5 | Q9aziz bacterioph |
| 10 | 40 | 54.1 | 542 | 16 Q9CJG6 | Q9cjpg6 lactococcus |
| 11 | 40 | 54.1 | 568 | 10 Q8RYG5 | Q8ryg5 oryza sativ |
| 12 | 40 | 54.1 | 594 | 16 Q9X156 | Q9x156 thermotoga |
| 13 | 39.5 | 53.4 | 1206 | 3 Q9C1F2 | Q9c1f2 pichia angu |
| 14 | 39 | 52.7 | 146 | 15 Q7ZRN4 | Q7zrn4 chimpanzee |
| 15 | 39 | 52.7 | 379 | 10 Q49391 | Q49391 arabidopsis |
| 16 | 39 | 52.7 | 437 | 16 Q9CL14 | Q9cll4 pasteurella |

| | | | | | |
|----|------|------|------|-----------|--------------------|
| 17 | 39 | 52.7 | 444 | 16 Q9HTW6 | Q9htw6 pseudomonas |
| 18 | 39 | 52.7 | 458 | 16 Q93N48 | Q93n48 coxiella bu |
| 19 | 39 | 52.7 | 530 | 16 Q67840 | Q67840 aquifex ao |
| 20 | 38.5 | 52.0 | 341 | 10 Q98RZ1 | Q98rz1 guillardia |
| 21 | 38.5 | 52.0 | 342 | 16 Q892M1 | Q892m1 bacteroides |
| 22 | 38 | 51.4 | 140 | 16 Q8XN55 | Q8xn55 clostridium |
| 23 | 38 | 51.4 | 317 | 16 Q899F1 | Q899f1 clostridium |
| 24 | 38 | 51.4 | 357 | 17 Q58509 | Q58509 pyrococcus |
| 25 | 38 | 51.4 | 358 | 12 Q9EMT0 | Q9emt0 ansacta moo |
| 26 | 38 | 51.4 | 362 | 3 Q9P3Y0 | Q9p3y0 zygosacchar |
| 27 | 38 | 51.4 | 390 | 10 Q9SW49 | Q9sw49 arabidopsis |
| 28 | 38 | 51.4 | 392 | 10 Q8LDM2 | Q8ldm2 arabidopsis |
| 29 | 38 | 51.4 | 563 | 16 Q8XJ35 | Q8xi35 clostridium |
| 30 | 38 | 51.4 | 965 | 10 Q8GT17 | Q8gt17 oryza sativ |
| 31 | 38 | 51.4 | 982 | 5 Q9BI28 | Q9bi28 cystosporog |
| 32 | 38 | 51.4 | 1034 | 10 Q7XW39 | Q7xw39 oryza sativ |
| 33 | 38 | 51.4 | 1045 | 16 Q8E808 | Q8e808 shewanella |
| 34 | 38 | 51.4 | 1645 | 10 Q7XS42 | Q7xs42 oryza sativ |
| 35 | 37.5 | 50.7 | 355 | 10 Q9SYM9 | Q9sym9 arabidopsis |
| 36 | 37 | 50.0 | 129 | 17 Q97XL6 | Q97xl6 sulfolobus |
| 37 | 37 | 50.0 | 135 | 4 Q8WYS0 | Q8wys0 homo sapien |
| 38 | 37 | 50.0 | 143 | 16 Q8PQ93 | Q8pq93 xanthomonas |
| 39 | 37 | 50.0 | 157 | 11 Q8K186 | Q8ki86 mus musculu |
| 40 | 37 | 50.0 | 161 | 16 Q82E14 | Q82e14 streptomyce |
| 41 | 37 | 50.0 | 266 | 16 Q97HM2 | Q97hm2 clostridium |
| 42 | 37 | 50.0 | 286 | 10 Q9SX55 | Q9sx55 arabidopsis |
| 43 | 37 | 50.0 | 298 | 5 Q9VAT6 | Q9vat6 drosophila |
| 44 | 37 | 50.0 | 323 | 12 Q9IGU9 | Q9igu9 tanapox vir |
| 45 | 37 | 50.0 | 323 | 12 Q9DHN9 | Q9dhn9 yaba-like d |

ALIGNMENTS

RESULT 1

Q8WWH6
ID Q8WWH6 PRELIMINARY; PRT; 316 AA.
AC Q8WWH6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Tumor antigen MAGE-N.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hepatoma;
RA Sui Y., Ye J., Wu W.;
RT "Cloning of a new gene of MAGE family in human hepatocellular carcinoma."
RT carcinoma."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF443295; AAL37897.1; -.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
SQ SEQUENCE 316 AA; 35409 MW; A463A9A740A089DF CRC64;

Query Match 100.0%; Score 74; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSARVRF 16
| | | | | | | | | | | | | | | |
Db 284 VKVLEYVIKVSARVRF 299

RESULT 2

Q14798
ID Q14798 PRELIMINARY; PRT; 317 AA.
AC Q14798;
DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MAGE-4 protein.
GN MELANOMA ANTIGEN-4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95369706; PubMed=7642112;
RA Imai Y., Shichiyo S., Yamada A., Katayama T., Yano H., Itoh K.;
RT "Sequence analysis of the MAGE gene family encoding human tumor-
rejection antigens.";
RL Gene 160:287-290(1995).
DR EMBL; D32075; BAA06841.1; -.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
SQ SEQUENCE 317 AA; 35044 MW; 9B9477253FE307C4 CRC64;

Query Match 75.7%; Score 56; DB 4; Length 317;
Best Local Similarity 73.3%; Pred. No. 0.068;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVR 15
Db |||||:|:|:|:|
285 VKVLEHVVRVNARVR 299

RESULT 3
Q9BUN9
ID Q9BUN9 PRELIMINARY; PRT; 318 AA.
AC Q9BUN9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Skin antigen, family A, 8 (Melanoma antigen, family A, 8).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor
vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002455; AAH02455.1; -.
DR EMBL; BC012744; AAH12744.1; -.
DR EMBL; BT007340; AAP36004.1; -.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
SQ SEQUENCE 318 AA; 35214 MW; EA02C1FB42F6C080 CRC64;

Query Match 75.7%; Score 56; DB 4; Length 318;
Best Local Similarity 73.3%; Pred. No. 0.068;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVR 15
Db |||||:|:|:|:|
287 VKVLEHVVRVNARVR 301

RESULT 4

Q8KEI5
ID Q8KEI5 PRELIMINARY; PRT; 248 AA.
AC Q8KEI5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc protease, putative.
GN CT0704.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AE012841; AAM71941.1; -.
DR TIGR; CT0704; -.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002725; DUF45.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01863; DUF45; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 248 AA; 28214 MW; 45382717E1A8A9F9 CRC64;

Query Match 56.8%; Score 42; DB 16; Length 248;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 LEYVIKVSARVRF 16
Db :||:|||||:
13 IEYTVKVSQRARY 25

RESULT 5
Q8ZZ01
ID Q8ZZ01 PRELIMINARY; PRT; 152 AA.
AC Q8ZZ01;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE0518.
GN PAE0518.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009771; AAL62840.1; -.
DR GO; GO:0008969; F:phosphohistidine phosphatase activity; IEA.
DR InterPro; IPR004449; SixA.
DR TIGRFAMs; TIGR00249; sixA; 1.

```

RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AE001812; AAD36806.1; -.
DR PIR; D72218; D72218.
DR TIGR; TM1741; -.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0008173; F:RNA methyltransferase activity; IEA.
DR GO; GO:0009451; P:RNA modification; IEA.
DR InterPro; IPR004441; rRNA_methyl_3.
DR InterPro; IPR001537; Spou_methylase.
DR Pfam; PF00588; Spou_methylase; 1.
DR ProDom; PD001243; Spou_methylase; 1.
DR TIGRFAMS; TIGR00186; rRNA_methyl_3; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 242 AA; 27096 MW; 03695F95E1921D14 CRC64;

Query Match 54.1%; Score 40; DB 16; Length 242;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSARVRF 16
Db 127 VKVTEIVVKVSAGTVF 142
||| | : ||| |
||| | : ||| |

RESULT 8
Q8WRT9 PRELIMINARY; PRT; 274 AA.
ID Q8WRT9;
AC Q8WRT9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Proteasome regulatory non-ATP-ase subunit 12.
GN RPN12.
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Z., Wang C.C.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF041120; AAL72635.1; -.
DR GO; GO:0005838; C:proteasome regulatory particle (sensu Eukarya); IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006746; Nln1_C.
DR Pfam; PF04653; Nln1_C; 1.
SQ SEQUENCE 274 AA; 31298 MW; 3FF0F1963953F83D CRC64;

Query Match 54.1%; Score 40; DB 5; Length 274;
Best Local Similarity 61.5%; Pred. No. 51;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KVLEYVIKVSARV 14
Db 69 EVLEYAVLVLSARL 81
: ||| | : ||| |
: ||| | : ||| |

RESULT 9
Q9AZI5 PRELIMINARY; PRT; 542 AA.
ID Q9AZI5
AC Q9AZI5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Helicase.
GN ORF24.
OS Bacteriophage bIL310.

```


OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=151538;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21113149; PubMed=11160885;
RA Chopin A., Bolotin A., Sorokin A., Ehrlich S.D., Chopin M.-C.;
RT "Analysis of six prophages in Lactococcus lactis IL1403: different
RT genetic structure of temperate and virulent phage populations.";
RL Nucleic Acids Res. 29:644-651(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Chopin A., Bolotin A., Sorokin A., Ehrlich S.D., Chopin M.-C.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF323671; AAK08428.1; -.
DR InterPro; IPR004968; Pox D5.
DR InterPro; IPR006500; Primase_C.
DR Pfam; PF03288; Pox D5; 1.
DR TIGRFAMs; TIGR01613; primase_Cterm; 1.
SQ SEQUENCE 542 AA; 62322 MW; 64CD85DE975CC5E4 CRC64;

Query Match 54.1%; Score 40; DB 9; Length 542;
Best Local Similarity 50.0%; Pred. No. 97;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 VLEYVIKVSARVRF 16
|||||:|:| :|
Db 402 VLEYVLKIALSLNF 415

RESULT 10
Q9CJG6 PRELIMINARY; PRT; 542 AA.
AC Q9CJG6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prophage psl protein 05, DNA primase.
GN Psl05 OR LL0029.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006242; AAK04127.1; -.
DR PIR; E86628; E86628.
DR InterPro; IPR004968; Pox D5.
DR InterPro; IPR006500; Primase_C.
DR Pfam; PF03288; Pox D5; 1.
DR TIGRFAMs; TIGR01613; primase_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 542 AA; 62322 MW; 64CD85DE975CC5E4 CRC64;

Query Match 54.1%; Score 40; DB 16; Length 542;
Best Local Similarity 50.0%; Pred. No. 97;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 VLEYVIKVSARVRF 16
|||||:|:| :|
Db 402 VLEYVLKIALSLNF 415

RESULT 11
Q8RYG5 PRELIMINARY; PRT; 568 AA.
ID Q8RYG5
AC Q8RYG5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Anther ethylene-upregulated protein-like.
GN P0592G05.12.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC
RT clone:P0592G05.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004672; BAB90826.1; -.
DR Gramene; Q8RYG5; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR000048; IQ_region.
DR Pfam; PF00023; ank; 2.
DR Pfam; PF00612; IQ; 3.
DR SMART; SM00248; ANK; 1.
DR SMART; SM00015; IQ; 3.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50096; IQ; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 568 AA; 64128 MW; 7E962344765200B3 CRC64;

Query Match 54.1%; Score 40; DB 10; Length 568;
Best Local Similarity 53.3%; Pred. No. 1e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VKVLEYVIKVSARVR 15
:|:|:|:|:|:|:|:|
Db 453 LKIRKNVIKIQARVR 467

RESULT 12
Q9X156 PRELIMINARY; PRT; 594 AA.
ID Q9X156
AC Q9X156;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein TM1334.
GN TM1334.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
DR EMBL; AE001788; AAD36406.1; -.
DR PIR; F72265; F72265.
DR TIGR; TM1334; -.
DR InterPro; IPR006158; B12-binding.
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF02310; B12-binding; 1.
DR Pfam; PF04055; Radical_SAM; 1.

DR SMART; SM00729; Elp3; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 594 AA; 68904 MW; 8D256B90F62BB07A CRC64;

Query Match 54.1%; Score 40; DB 16; Length 594;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 LEYVIKVSARVR 15
|||:||||
Db 288 LEYIKKGRVR 299

RESULT 13
Q9C1F2 PRELIMINARY; PRT; 1206 AA.
AC Q9C1F2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Alpha-COP-like protein.
GN OPU27.
OS Pichia angusta (Yeast) (Hansenula polymorpha).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4905;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS4732;
RA Agaphonov M.O.; Chechenova M.B.; Deev A.V.; Packeiser A.N.;
RA Ter-Avanesyan M.D.;
RT "Hansenula polymorpha alpha-COP homolog."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF339501; AAK26326.1; -.
DR InterPro; IPR006692; Coatomer_WDAD.
DR InterPro; IPR001680; WD40.
DR Pfam; PF04053; Coatomer_WDAD; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 5.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 1206 AA; 135629 MW; E6F1212657DBEE14 CRC64;

Query Match 53.4%; Score 39.5; DB 3; Length 1206;
Best Local Similarity 57.9%; Pred. No. 2.5e+02;
Matches 11; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

Qy 1 VKVLE---YVIKVSARVRF 16
:|||||
Db 572 IKTLENTLYVIKVSGRQCF 590

RESULT 14
Q7ZRN4 PRELIMINARY; PRT; 146 AA.
AC Q7ZRN4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SMM.US.-.M922;
RA Ling B.; Santiago M.L.; Meleth S.; Gormus B.; McClure H.M.;
RA Apetrei C.; Hahn B.H.; Marx P.A.;

RT "Non-Invasive Detection of New Simian Immunodeficiency Virus Lineages
in Captive Sooty Mangabays: Ability to Amplify Virion RNA from Fecal
Samples Correlates with Plasma Viral Load";
RL J. Virol. 0:0-0(2003).
DR EMBL; AY159611; AAO18469.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
FT NON_TER 1
FT NON_TER 146 146
SQ SEQUENCE 146 AA; 17398 MW; 329A10B9B2435486 CRC64;

Query Match 52.7%; Score 39; DB 15; Length 146;
Best Local Similarity 40.0%; Pred. No. 43;
Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKVSARVR 15
:::|:::|
Db 129 LRIVIVVQMSARLR 143

RESULT 15
O49391 PRELIMINARY; PRT; 379 AA.
AC O49391;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine/threonine protein kinase like protein.
GN F10N7.190 OR AT4G32000.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M.; Koetter P.; Hempel S.; Entian K.-D.; Hoheisel J.;
RA Mewes H.W.; Mayer K.F.X.; Schueller C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Terryn N.; Ardiles W.; Buysshaert C.; Dasseville R.; De Clerck R.;
RA De Keyser A.; Neyt P.; Rouze P.; Van Den Daele H.; Villarroel R.;
RA Gielen J.; Van Montagu M.; Mewes H.W.; Lemcke K.; Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Rose M.; Hempel S.; Entian K.-D.; Mewes H.W.; Lemcke K.; Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL021636; CAA16589.1; -.
DR EMBL; AL161580; CAB79918.1; -.
DR PIR; T04645; T04645.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 2.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 379 AA; 42047 MW; 80750DE264E6E0FB CRC64;

Query Match 52.7%; Score 39; DB 10; Length 379;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VKVLEYVIKUSA 12
| | | | : | | |
Db 14 VSVLEFIISVSA 25

Search completed: October 7, 2004, 15:16:57
Job time : 60.8182 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 12:53:09 ; Search time 62.9091 Seconds
(without alignments)
71.862 Million cell updates/sec

Title: US-09-336-091-4
Perfect score: 80
Sequence: 1 EYVIKVSARVFFFPs 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 80 | 100.0 | 16 | 4 AAB31292 | Aab31292 Peptide d |
| 2 | 80 | 100.0 | 30 | 5 AAU85042 | Aau85042 Human MAG |
| 3 | 80 | 100.0 | 309 | 2 AAR70909 | Aar70909 Human mel |
| 4 | 80 | 100.0 | 309 | 2 AAW81548 | Aaw81548 Tumour re |
| 5 | 80 | 100.0 | 309 | 4 AAB31290 | Aab31290 Amino aci |
| 6 | 80 | 100.0 | 309 | 4 AAE06806 | Aae06806 Human MAG |
| 7 | 80 | 100.0 | 309 | 5 AAU84814 | Aau84814 Human MAG |
| 8 | 80 | 100.0 | 309 | 6 ABP74195 | Abp74195 Human MAG |
| 9 | 80 | 100.0 | 309 | 6 ABU08930 | Abu08930 Human tum |
| 10 | 80 | 100.0 | 309 | 7 ADC09573 | Adc09573 MAGE-1 pr |
| 11 | 80 | 100.0 | 310 | 6 AAO19742 | Aao19742 Wild-type |
| 12 | 80 | 100.0 | 316 | 6 ABU04419 | Abu04419 Human exp |
| 13 | 80 | 100.0 | 445 | 2 AAY06592 | Aay06592 CLYTA-MAG |
| 14 | 80 | 100.0 | 446 | 2 AAY06590 | Aay06590 Lipoprote |
| 15 | 80 | 100.0 | 1052 | 6 ABR57354 | AbR57354 MatDC16-C |
| 16 | 80 | 100.0 | 3541 | 5 AAU85130 | Aau85130 Human mel |
| 17 | 64 | 80.0 | 28 | 5 AAU85043 | Aau85043 Human MAG |
| 18 | 60 | 75.0 | 12 | 4 AAB31296 | Aab31296 Peptide d |
| 19 | 57 | 71.2 | 12 | 4 AAB31295 | Aab31295 Peptide d |
| 20 | 57 | 71.2 | 16 | 4 AAB31291 | Aab31291 Peptide d |
| 21 | 51 | 63.7 | 317 | 2 AAY06998 | Aay06998 MAGE-4 pr |
| 22 | 51 | 63.7 | 317 | 6 ABU56545 | Abu56545 Lung canc |
| 23 | 51 | 63.7 | 317 | 6 ADA83770 | Ada83770 Human MAG |
| 24 | 51 | 63.7 | 318 | 3 AAB08734 | Aab08734 Amino aci |
| 25 | 49 | 61.3 | 10 | 2 AAR70951 | Aar70951 Human mel |

| | | | | | |
|----|----|------|----|------------|--------------------|
| 26 | 49 | 61.3 | 10 | 2 AAY46180 | Aay46180 Immunogen |
| 27 | 47 | 58.8 | 16 | 4 AAB31301 | Aab31301 Peptide d |
| 28 | 46 | 57.5 | 10 | 2 AAY38326 | Aay38326 MAGE-deri |
| 29 | 46 | 57.5 | 10 | 2 AAR70948 | Aar70948 Human mel |
| 30 | 46 | 57.5 | 10 | 2 AAR70960 | Aar70960 Human mel |
| 31 | 46 | 57.5 | 10 | 2 AAY46118 | Aay46118 Immunogen |
| 32 | 46 | 57.5 | 10 | 2 AAY45907 | Aay45907 Immunogen |
| 33 | 46 | 57.5 | 10 | 6 ABP74245 | Abp74245 Human MAG |
| 34 | 46 | 57.5 | 10 | 7 ADC09104 | Adc09104 Epitope w |
| 35 | 45 | 56.2 | 10 | 2 AAR70953 | Aar70953 Human mel |
| 36 | 45 | 56.2 | 10 | 2 AAY46179 | Aay46179 Immunogen |
| 37 | 44 | 55.0 | 9 | 2 AAR70933 | Aar70933 Human mel |
| 38 | 44 | 55.0 | 9 | 2 AAY46137 | Aay46137 Immunogen |
| 39 | 43 | 53.8 | 9 | 2 AAR70927 | Aar70927 Human mel |
| 40 | 43 | 53.8 | 9 | 2 AAR70921 | Aar70921 Human mel |
| 41 | 43 | 53.8 | 9 | 2 AAY46136 | Aay46136 Immunogen |
| 42 | 43 | 53.8 | 9 | 2 AAY47262 | Aay47262 Immunogen |
| 43 | 43 | 53.8 | 9 | 3 AAY92316 | Aay92316 MAGE-A1 a |
| 44 | 43 | 53.8 | 9 | 3 AAY92274 | Aay92274 MAGE-A1 a |
| 45 | 43 | 53.8 | 9 | 5 AAE31182 | Aae31182 Human mag |

ALIGNMENTS

RESULT 1
AAB31292
ID AAB31292 standard; peptide; 16 AA.
XX
AC AAB31292;
XX
DT 20-APR-2001 (first entry)
XX
DE Peptide derived from human MAGE-A1 HLA class II-binding protein.
XX
KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;
KW MAGE-A1 HLA class II-binding protein; vaccine.
XX
OS Homo sapiens.
XX
PN WO200078806-A1.
XX
PD 28-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US016287.
XX
PR 18-JUN-1999; 99US-00336091.
XX
(LUDW-) LUDWIG INST CANCER RES.
PI Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;
XX
WPI; 2001-102698/11.
XX
PT Novel MAGE-A1 human leukocyte antigen class II peptides which bind to and
PT are presented to the class II molecules, useful for inducing immune
PT response and treating cancers characterized by expression of MAGE-A1.
XX
PS Claim 3; Page 42; 78pp; English.
XX
CC The present sequence is derived from a human MAGE-A1 HLA (human leukocyte
CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA
CC binding protein stimulate the activity and proliferation of CD4+ T
CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic
CC agent for diagnosing a disorder characterized by expression of MAGE-A1.
CC The protein is used for treating a disorder characterized by expression
CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,
CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias. Peptides
CC derived from the MAGE-A1 HLA binding protein are useful in the production
XX of anti-tumour vaccines
SQ Sequence 16 AA;

Query Match 100.0%; Score 80; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 6.5e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFFPSS 16
| | | | | | | | | | | | | | | |
Db 1 EYVIKVSARVRRFFFPSS 16

RESULT 2
AAU85042
ID AAU85042 standard; peptide; 30 AA.
XX
AC AAU85042;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human MAGE-1 segment 19.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Homo sapiens.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
DR N-PSDB; ABK36862.
XX
PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.
XX
PS Example 3; Fig 27; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
CC a peptide derived from a parent protein used to construct a savine of the
CC invention
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 80; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFFPSS 16
| | | | | | | | | | | | | | | |
Db 13 EYVIKVSARVRRFFFPSS 28

RESULT 3
AAR70909
ID AAR70909 standard; protein; 309 AA.
XX
AC AAR70909;
XX
DT 25-MAR-2003 (revised)
DT 09-OCT-1995 (first entry)
XX
DE Human melanoma antigen MAGE-1.
XX
KW Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW HLA-restricted cytotoxic T-lymphocyte activity.
XX
OS Homo sapiens.
XX
PN WO9504542-A1.
XX
PD 16-FEB-1995.
XX
PF 02-AUG-1994; 94WO-US008721.
XX
PR 06-AUG-1993; 93US-00103623.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Fikes JD, Livingston BD, Sette AD, Sidney JC;
XX
DR WPI; 1995-090681/12.
DR N-PSDB; AAQ85435.
XX
PT Human melanoma antigen, MAGE-1, peptide(s) - useful for stimulating
PT immune response against melanoma.
XX
PS Example1; Fig 1; 59pp; English.
XX
CC AAQ85435 encodes AAR70909 human melanoma antigen MAGE-1, it was used to
CC produce the C-terminal MAGE-1 peptides described in AAR70915 to AAR70969.
CC These peptides are useful for defining epitopes that engender a HLA-
CC restricted cytotoxic lymphocyte activity against MAGE-1 antigens.
CC Compsns. containing these peptides can be administered, as a vaccine to
CC patients susceptible to MAGE associated tumours, e.g. melanomas. (Updated
CC on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 80; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFFPSS 16
| | | | | | | | | | | | | | | |
Db 281 EYVIKVSARVRRFFFPSS 296

RESULT 4
AAW81548
ID AAW81548 standard; protein; 309 AA.
XX
AC AAW81548;
XX
DT 01-MAR-1999 (first entry)
XX

DE Tumour rejection antigen precursor MAGE-A1.
XX MAGE-A1; human; tumour rejection antigen precursor; TRAP; therapy;
KW diagnosis.
XX Homo sapiens.
OS
XX
PN WO9849184-A1.
XX
XX
PD 05-NOV-1998.
XX
XX 24-APR-1998; 98WO-US008493.
XX
PR 25-APR-1997; 97US-00845528.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
PI Lucas S, De Smet C, Boon-Falleur T;
XX
DR WPI; 1999-024041/02.
DR N-PSDB; AAV69719.
XX
PT Tumour rejection antigen precursors - used for determining presence of
PT cytolytic T cells specific for complexes of a human leukocyte antigen.
XX
PS Disclosure; Page 50-51; 84pp; English.
XX
CC This is the amino acid sequence of human tumour rejection antigen
CC precursor (TRAP) MAGE-A1. MAGE-A1 cDNA (see AAV69719) shows homology to
CC novel human MAGE-C1 cDNA (see AAV69720). MAGE-C1 (see AAW81546) is a
CC novel member of the MAGE family that may be recognised by cytotoxic T
CC cells, leading to lysis of the tumour cells which express it. It is
CC expressed in a variety of tumours and in normal testis cells, but not by
CC other normal cells. The invention provides MACE-C1 and MACE-C2 nucleic
CC acids and polypeptides, useful e.g. in a claimed method for determining
CC the presence of cytolytic T cells specific for complexes of a human
CC leukocyte antigen (HLA)
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 80; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFPS 16
Db 281 EYVIKVSARVRRFFPS 296
|||||

RESULT 5
AAB31290
ID AAB31290 standard; protein; 309 AA.
XX
AC AAB31290;
XX
DT 20-APR-2001 (first entry)
XX
DE Amino acid sequence of human MAGE-A1 HLA class II-binding protein.
XX
KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;
KW MAGE-A1 HLA class II-binding protein; vaccine.
XX
XX Homo sapiens.
OS
XX
PN WO200078806-A1.
XX
PD 28-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US016287.
XX
PR 18-JUN-1999; 99US-00336091.
XX
PA (LUDW-) LUDWIG INST CANCER RES.

XX Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;
PI WPI; 2001-102698/11.
XX N-PSDB; AAF24676.
DR
XX
XX Novel MAGE-A1 human leukocyte antigen class II peptides which bind to and
PT are presented to the class II molecules, useful for inducing immune
PT response and treating cancers characterized by expression of MAGE-A1.
XX
XX Claim 1; Page 63; 78pp; English.
PS
XX
CC The present sequence represents a human MAGE-A1 HLA (human leukocyte
CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA
CC binding protein stimulate the activity and proliferation of CD4+ T
CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic
CC agent for diagnosing a disorder characterized by expression of MAGE-A1.
CC The protein is used for treating a disorder characterized by expression
CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,
CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias. Peptides
CC derived from the MAGE-A1 HLA binding protein are useful in the production
CC of anti-tumour vaccines
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 80; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFPS 16
Db 281 EYVIKVSARVRRFFPS 296
|||||

RESULT 6
AAE06806
ID AAE06806 standard; protein; 309 AA.
XX
AC AAE06806;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human MAGE-A1 protein.
XX
KW MAGE antigenic peptide; Human leukocyte antigen; HLA-B35; HLA-B44;
KW tumour cell; immunostimulant; antigen presentation; cancer; melanoma;
KW CD8+ cytotoxic T lymphocyte; colorectal; prostate; gastric carcinoma;
KW myeloma; brain tumour; sarcoma; seminoma; ovarian tumour; cytostatic;
KW gene therapy; human; MAGE-A1; tumour rejection antigen; TRA.
XX
OS Homo sapiens.
XX
PN WO200153833-A1.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US002008.
XX
PR 20-JAN-2000; 2000US-0177242P.
PR 25-OCT-2000; 2000US-0243212P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Luiten R, Boon-Falleur T, Van Der Bruggen P, Stroobant V;
PI Demotte N, Schultz E;
XX
DR WPI; 2001-488724/53.
DR N-PSDB; AAD12987.
XX
PT Functional variants and isolated mimetics of a MAGE-A1 HLA-B35 or HLA-B44
PT binding peptide, or of a MAGE-A3 HLA-B35 binding peptide, used in
PT diagnosis and treatment of a disorder characterized by expression of MAGE
PT -A1 or -A3.

XX Claim 2; Page 86-87; 103pp; English.

PS The invention relates to functional variants and isolated mimetics of a

XX MAGE-A1 human leukocyte antigen (HLA)-B35 or HLA-B44 binding peptide, or

CC of a MAGE-A3 HLA-B35 binding peptide, identified by methods described in

CC the specification. MAGE genes encode tumour rejection antigens (TRAs)

CC presented to T lymphocytes by HLA-B35 and HLA-B44 molecules. The MAGE

CC antigenic peptide acts by binding to HLA molecules on tumour cells and

CC stimulating recognition of these cells and thus signalling them to the

CC immune system for destruction. The peptide when presented by HLA molecule

CC induces the activation and stimulation of CD8+ cytotoxic T lymphocytes.

CC The MAGE antigenic peptide is used to treat and diagnose disorders

CC characterised by expression of MAGE-A1 or -A3. Disorders include cancers

CC e.g melanomas, oesophageal, lung, head and neck, breast, colorectal,

CC prostate, renal, bladder, hepatocellular, papillary thyroid and gastric

CC carcinomas, myelomas, brain tumours, sarcomas, seminomas, and ovarian

CC tumours. The present sequence is human MAGE-A1 protein

XX

SQ Sequence 309 AA;

Query Match 100.0%; Score 80; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. NO. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 281 EYVIKVSARVRFPPS 296

RESULT 7

AAU84814

ID AAU84814 standard; protein; 309 AA.

XX

AC AAU84814;

XX

DT 08-MAY-2002 (first entry)

XX

DE Human MAGE-1 consensus sequence.

XX

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

KW viral infection; human immunodeficiency virus; melanoma;

KW bacterial infection; Salmonella; Legionella; parasitic infection;

KW Trypanosoma; Toxoplasma; Giardia.

XX

OS Homo sapiens.

XX

PN WO200190197-A1.

XX

PD 29-NOV-2001.

XX

PF 25-MAY-2001; 2001WO-AU000622.

XX

PR 26-MAY-2000; 2000AU-00007761.

XX

PA (AUSU) UNIV AUSTRALIAN NAT.

XX

PI Thomson SA, Ramshaw IA;

XX

DR WPI; 2002-147575/19.

XX

PT New synthetic polypeptides having several different segments of at least

PT one parent polypeptide linked together differently compared to the

PT linkage in the parent polypeptide, for inducing immune response against a

PT pathogen or cancer.

XX

PS Example 3; Fig 27; 364pp; English.

XX

CC The invention relates to a new synthetic polypeptide (I) comprising

CC several different segments of at least one parent polypeptide linked

CC together in a different relationship relative to their linkage in the

CC parent polypeptide to impede, abrogate or otherwise alter at least one

CC function associated with the parent polypeptide and for inducing an

CC immune response against a pathogen or cancer. Also included are a

CC synthetic polynucleotide encoding and a computer system for designing the

CC synthetic polypeptides. The synthetic polypeptides and polynucleotides

CC are referred to as a Savine. The synthetic polypeptide is useful for

CC modulating immune responses preferably directed against a pathogen or a

CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head

CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,

CC oesophagus, brain, testicle, uterus), as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or

CC prophylaxis against viral (such as infections caused by HIV (human

CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis

CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial

CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,

CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic

CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,

CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is

CC a consensus sequence for a parent protein used to design a savine of the

CC invention

XX

SQ Sequence 309 AA;

Query Match 100.0%; Score 80; DB 5; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 281 EYVIKVSARVRFPPS 296

RESULT 8

ABP74195

ID ABP74195 standard; protein; 309 AA.

XX

AC ABP74195;

XX

DT 03-FEB-2003 (first entry)

XX

DE Human MAGE-1 protein SEQ ID NO:71.

XX

KW Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;

KW T cell.

XX

OS Homo sapiens.

XX

PN WO200281646-A2.

XX

PD 17-OCT-2002.

XX

PF 04-APR-2002; 2002WO-US011101.

XX

PR 06-APR-2001; 2001US-0282211P.

PR 07-NOV-2001; 2001US-0337017P.

PR 07-MAR-2002; 2002US-0363210P.

XX

PA (CTLI-) CTL IMMUNOTHERAPIES CORP.

XX

PI Simard JTL, Diamond DC, Liu L, Xie Z;

XX

DR WPI; 2003-067518/06.

DR N-PSDB; ABQ83847.

XX

PT Novel epitopes useful as vaccines, comprises peptides or nucleic acid

PT encoding the peptides, that are useful epitopes of target-associated

PT antigens.

XX

PS Claim 1; Page 156; 352pp; English.

XX

CC The present invention describes an isolated epitope (I) and an epitope

CC cluster. Also described is a vaccine or immunotherapeutic composition

CC (VC) comprising (I). (I) has cytostatic activity. VC is useful for

CC treating an animal, by administering to an animal the vaccine or

CC immunotherapeutic composition. VC is also useful for evaluating

CC immunogenicity of a vaccine or immunotherapeutic composition, by
CC administering VC to an HLA-transgenic animal and evaluating
CC immunogenicity based on a characteristic of the animal, or by in vitro
CC primary stimulation of a T cell and evaluating immunogenicity. (I) is
CC useful for determining specific T cell frequency, by contacting T cells
CC with a MHC-peptide complex, and further comprises ELISPOT analysis,
CC limiting dilution analysis, flow cytometry, in situ hybridisation and/or
CC polymerase chain reaction (PCR). ABQ83843 to ABQ83858 and ABP74128 to
CC ABP74713 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 80; DB 6; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVRFPPS 16
Db 281 EYVIKVSARVRFPPS 296

RESULT 9
ABU08930
ID ABU08930 standard; protein; 309 AA.
XX
AC ABU08930;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human tumour rejection antigen precursor, MAGE-A1.
XX
KW TRAP; tumour rejection antigen precursor; cytolytic T-cell; CTL; tumour;
KW seminoma; bladder transitional-cell carcinoma; NSCLC; adaptor;
KW head-and-neck squamous-cell carcinoma; breast carcinoma; sarcoma;
KW cutaneous melanoma; non-small cell lung cancer; MAGE-A1; human.
XX
OS Homo sapiens.
XX
PN US2002176865-A1.
XX
PD 28-NOV-2002.
XX
PF 01-MAR-2002; 2002US-00085108.
XX
PR 25-APR-1997; 97US-00845528.
PR 24-APR-1998; 98US-00066281.
PR 17-DEC-1999; 99US-00468433.
PR 09-FEB-2000; 2000US-00501104.
XX
PA (LUCA/) LUCAS S.
PA (BOON/) BOON-FALLEUR T.
XX
PI Lucas S, Boon-Falleur T;
XX
DR WPI; 2003-328468/31.
DR N-PSDB; ABX93696.
XX
PT Novel isolated nucleic acid encoding tumor rejection antigen precursor
PT MAGE-C3, MAGE-B5, or MAGE-B6, useful as diagnostic probes to determine
PT presence of abnormal e.g., tumor cells expressing MAGE-C1, MAGE-B5 or
PT MAGE-B6.
XX
PS Disclosure; Fig 2; 59pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule which encodes
CC a tumour rejection antigen precursor (TRAP) having an amino acid sequence
CC of a TRAP encoded by a fully defined MAGE-C3, MAGE-B5, or MAGE-B6
CC polynucleotide sequence. Also disclosed is a method which is useful for
CC determining presence of cytolytic T-cells specific for complexes of human
CC leukocyte antigen (HLA) and a peptide derived from the nucleic acid in a
CC cytotoxic T-lymphocyte (CTL)-containing sample. The nucleic acid is
CC useful as a diagnostic probe to determine the presence of abnormal

CC (tumour) cells such as seminoma, bladder transitional-cell carcinoma,
CC head-and-neck squamous-cell carcinoma, breast carcinoma, sarcoma,
CC cutaneous melanoma or non-small cell lung cancer (NSCLC) which express
CC MAGE-C1, MAGE-B5 or MAGE-B6. The nucleic acid is useful for diagnosing a
CC disorder characterised by expression of MAGE-C1, MAGE-B5 or MAGE-B6 TRAPs
CC or tumour rejection antigens (TRAs). The present sequence represents the
CC amino acid sequence of the human tumour rejection antigen precursor, MAGE
CC -A1
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 80; DB 6; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVRFPPS 16
Db 281 EYVIKVSARVRFPPS 296

RESULT 10
ADC09573
ID ADC09573 standard; protein; 309 AA.
XX
AC ADC09573;
XX
DT 18-DEC-2003 (first entry)
XX
DE MAGE-1 protein #SEQ ID 71.
XX
KW Epitope; immunological; vaccine;
KW major histocompatibility complex class I; MHC class I; cancer;
KW immunisation.
XX
OS Unidentified.
XX
PN WO2003008537-A2.
XX
PD 30-JAN-2003.
XX
PF 29-MAR-2002; 2002WO-US010189.
XX
PR 06-APR-2001; 2001US-0282211P.
PR 07-NOV-2001; 2001US-0337017P.
PR 07-MAR-2002; 2002US-0363210P.
XX
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX
PI Simard JJJL, Diamond DC, Liu L, Xie Z;
XX
DR WPI; 2003-248010/24.
XX
PT Epitope having high affinity for major histocompatibility complex class I
PT useful for treating an animal, evaluating immunogenicity of a vaccine or
PT therapeutic composition and for diagnosing a disease.
XX
PS Claim 1; SEQ ID NO 71; 239pp; English.
XX
CC The invention relates to an isolated epitope polypeptide that has high
CC affinity for major histocompatibility complex (MHC) class I, and an
CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine
CC or immunotherapeutic composition containing an epitope of the invention.
CC Compositions of the invention may be used in the treatment of cancer. The
CC method can be combined with a radiation therapy, chemotherapy,
CC biochemotherapy or surgery. The composition is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
CC -peptide complexes of the invention are useful for determining specific T
CC cell frequency. This method is useful for evaluating immunological
CC response, by performing the method prior to and subsequent to an
CC immunisation step. Compositions of the invention are useful for
CC diagnosing a disease. The current sequence represents an epitope of the
CC invention with high affinity for MHC class I.
XX

SQ Sequence 309 AA;
Query Match 100.0%; Score 80; DB 7; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EYVIKVSARVRRFFPS 16
| | | | | | | | | | | | | | | |
Db 281 EYVIKVSARVRRFFPS 296
RESULT 11
AAO19742
ID AAO19742 standard; protein; 310 AA.
XX
AC AAO19742;
XX
DT 11-AUG-2003 (first entry)
XX
DE Wild-type MAGE1 protein.
XX
KW Stabilised mRNA; translation optimised; vaccine; tissue repair;
KW sequence modification determination; gene therapy; cytostatic; virucide;
KW antibacterial; protozoacide; nootropic; neuroprotective; infection;
KW antiparkinsonian; immunostimulant; cancer; MAGE1 protein.
XX
OS Unidentified.
XX
PN WO200298443-A2.
XX
PD 12-DEC-2002.
XX
PF 05-JUN-2002; 2002WO-EP006180.
XX
PR 05-JUN-2001; 2001DE-01027283.
XX
PA (VMUE/) VON DER MUELBE F.
XX
PI Von Der Muelbe F, Hoerr I, Pascolo S;
XX
DR WPI; 2003-148621/14.
DR N-PSDB; ABZ69107.
XX
PT Composition containing mRNA modified for optimal translation and
PT stability, useful for treating e.g. tumors or infections, comprises
PT increased G/C content and fewer rare codons.
XX
PS Disclosure; Fig 2B; 75pp; German.
XX
CC The present invention relates to a pharmaceutical composition containing
CC at least one modified RNA encoding a biologically active or antigenic
CC protein. The RNA is modified to optimise translation of the sequence. The
CC compositions are used for vaccination against a wide range of infectious
CC diseases (viral, bacterial or protozoal) or cancer, or for tissue
CC regeneration, e.g. in cases of Alzheimer's or Parkinson's diseases and
CC arthritis, but also to express proteins such as dystrophins, chloride ion
CC channels (for treating cystic fibrosis) and enzymes (either for treating
CC metabolic disorders or for synthesis of neurotransmitters such as
CC dopamine). The present sequence is the wild-type MAGE1 protein
XX
SQ Sequence 310 AA;
Query Match 100.0%; Score 80; DB 6; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EYVIKVSARVRRFFPS 16
| | | | | | | | | | | | | | | |
Db 280 EYVIKVSARVRRFFPS 295
RESULT 12
ABU04419

ID ABU04419 standard; protein; 316 AA.
XX
AC ABU04419;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1085.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicz RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 1085; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 316 AA;
Query Match 100.0%; Score 80; DB 6; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EYVIKVSARVRRFFPS 16
| | | | | | | | | | | | | | | |
Db 288 EYVIKVSARVRRFFPS 303
RESULT 13
AAV06592

ID AAY06592 standard; protein; 445 AA.
XX
AC AAY06592;
XX
DT 26-OCT-1999 (first entry)
XX
DE CLYTA-MAGE-1-His fusion protein.
XX
KW MAGE-1; CLYTA-MAGE-1-His; fusion protein; tumour; melanoma;
KW breast cancer; bladder cancer; lung cancer; colon cancer;
KW head and squamous cell carcinoma; oesophagus carcinoma; vaccine; human.
XX
OS Streptococcus pneumoniae.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX WO9940188-A2.
PN
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-EP0000660.
XX
PR 05-FEB-1998; 98GB-00002543.
PR 06-FEB-1998; 98GB-00002650.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cabezon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;
XX
DR WPI; 1999-494293/41.
DR N-PSDB; AAX87593.
XX
PT New protein derivatives used in cancer vaccine therapy for treating a
PT range of cancers including melanomas, carcinomas and cancers of breast.
XX
PS Example 9; Page 69-70; 72pp; English.
XX
CC The present sequence represents a fusion protein composed of the C-
CC terminal portion of the Streptococcus pneumoniae LYTA protein (CLYTA),
CC the human MAGE-1 tumour-associated antigen and a hexahistidine tail. A
CC vector designed for recombinant expression of the fusion protein in
CC Escherichia coli is provided. The CLYTA moiety provides expression of
CC soluble fusion protein, facilitates affinity purification, and also acts
CC as a T-helper epitope. The invention relates to MAGE proteins fused to an
CC immunological fusion partner, e.g. CLYTA-MAGE-1-His. These novel fusion
CC proteins provide vaccines for immunotherapy of melanomas or other MAGE-
CC associated tumours like breast, bladder, lung and non-small cell lung
CC cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus
CC carcinoma
XX
SQ Sequence 445 AA;
Query Match 100.0%; Score 80; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 408 EYVIKVSARVRFPPS 423
RESULT 14
AAY06590
ID AAY06590 standard; protein; 446 AA.
XX
AC AAY06590;
XX
DT 26-OCT-1999 (first entry)
XX
DE Lipoprotein D-MAGE-1-His fusion protein.
XX
KW MAGE-1; lipoprotein D; LPD-MAGE-1-His; fusion protein; tumour; melanoma;

KW breast cancer; bladder cancer; lung cancer;
KW head and squamous cell carcinoma; colon cancer; oesophagus carcinoma;
KW vaccine; human.
XX
OS Haemophilus influenzae.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX WO9940188-A2.
PN
XX 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-EP0000660.
XX
PR 05-FEB-1998; 98GB-00002543.
PR 06-FEB-1998; 98GB-00002650.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cabezon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;
XX
DR WPI; 1999-494293/41.
DR N-PSDB; AAX87591.
XX
PT New protein derivatives used in cancer vaccine therapy for treating a
PT range of cancers including melanomas, carcinomas and cancers of breast.
XX
PS Example 6; Page 67-68; 72pp; English.
XX
CC The present sequence represents a novel fusion protein composed of
CC lipidated protein D (LPD) of Haemophilus influenzae B, the human MAGE-1
CC tumour-associated antigen and a hexahistidine tail. The invention relates
CC to MAGE proteins fused to an immunological fusion partner such as LPD.
CC The LPD moiety provides the fusion protein with additional exogenous T-
CC cell epitopes and also increase expression levels in E. coli. The lipid
CC tail ensures optimal presentation of the antigen to antigen-presenting
CC cells. The affinity tag facilitates purification. The novel fusion
CC proteins provide vaccines for immunotherapy of melanomas or other MAGE-
CC associated tumours like breast, bladder, lung and non-small cell lung
CC cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus
CC carcinoma
XX
SQ Sequence 446 AA;
Query Match 100.0%; Score 80; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 409 EYVIKVSARVRFPPS 424
RESULT 15
ABR57354
ID ABR57354 standard; protein; 1052 AA.
XX
AC ABR57354;
XX
DT 09-SEP-2003 (first entry)
XX
DE MatDC16-C-gamma-4-MAGE-A1 amino acid sequence.
XX
KW Antigen presenting cell; vaccination; neurotropic; neuroprotective;
KW antiarteriosclerotic; cytostatic; antidiabetic; hepatotropic;
KW antiinflammatory; antiparasitic; fungicide; antibacterial; virucide;
KW vaccine; Alzheimer's disease; atherosclerosis; cancer; diabetes;
KW hepatitis; infection.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers

FT Misc-difference 546 /note= "unspecified"

XX WO2003046011-A1.

XX PD 05-JUN-2003.

XX PF 30-NOV-2001; 2001WO-EP014255.

XX PR 30-NOV-2001; 2001WO-EP014255.

XX PA (CRUC-) CRUCELL HOLLAND BV.

XX PI Germeraad W;

XX DR WPI; 2003-493401/46.

XX PT New conjugate for targeting antigen presenting cells, useful for
PT preventing, retarding or treating e.g., Alzheimer's disease,
PT atherosclerosis, cancer, diabetes, hepatitis or fungal, bacterial or
PT viral infections.

XX PS Disclosure; Fig 2; 54pp; English.

XX CC The present invention describes a conjugate (I) for targeting antigen
CC presenting cells (APCs) comprising at least one antigenic moiety
CC conjugated to a targeting moiety that is capable of binding to a cell
CC surface structure of an APC, and upon binding, inducing a cytotoxic T
CC lymphocyte (CTL) and T-helper response. Also described: (1) a nucleic
CC acid sequence encoding the antigenic or targeting moiety; (2) an
CC expression vector comprising the nucleic acid sequence, operably linked
CC to expression sequences for the APC; (3) a host cell transformed or
CC transfected using the nucleic acid or expression vector; (4) a method for
CC producing (I); (5) a method for generating an APC, capable of eliciting
CC an immune response via MHC classes I and II presentation of processed
CC antigen fragments; and (6) a pharmaceutical composition comprising (I) or
CC the APC. (I) has nootropic, neuroprotective, virucide,
CC antiarteriosclerotic, cytostatic, antidiabetic, hepatotropic, fungicide,
CC antiinflammatory, antiparasitic and antibacterial activities, and can be
CC used in vaccines. The conjugate (I) or APC can be used for preventing,
CC retarding or treating e.g., Alzheimer's disease, atherosclerosis,
CC cancer, diabetes, hepatitis or parasitic, fungal, bacterial or viral
CC infections. The present sequence represents a MatDC16-C-gamma-4-MAGE-A1
CC amino acid sequence, which is used in the exemplification of the present
CC invention

XX SQ Sequence 1052 AA;

Query Match 100.0%; Score 80; DB 6; Length 1052;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVFFPPS 16

Db 1024 EYVIKVSARVFFPPS 1039

Search completed: October 7, 2004, 15:13:02
Job time : 63.9091 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:18:26 ; Search time 17.4545 Seconds
(without alignments)
47.324 Million cell updates/sec

Title: US-09-336-091-4
Perfect score: 80
Sequence: 1 EYVIKVSARVRRFFFPS 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 80 | 100.0 | 58 | 1 US-08-465-167A-1 | Sequence 1, Appli |
| 2 | 80 | 100.0 | 58 | 4 US-08-627-820-1 | Sequence 1, Appli |
| 3 | 80 | 100.0 | 309 | 1 US-08-465-167A-24 | Sequence 24, Appl |
| 4 | 80 | 100.0 | 309 | 2 US-08-993-118-10 | Sequence 10, Appl |
| 5 | 80 | 100.0 | 309 | 3 US-08-845-528C-10 | Sequence 10, Appl |
| 6 | 80 | 100.0 | 309 | 4 US-08-627-820-24 | Sequence 24, Appl |
| 7 | 80 | 100.0 | 309 | 4 US-09-066-281B-10 | Sequence 10, Appl |
| 8 | 80 | 100.0 | 309 | 4 US-09-468-433C-10 | Sequence 10, Appl |
| 9 | 80 | 100.0 | 309 | 4 US-09-392-714-29 | Sequence 29, Appl |
| 10 | 49 | 61.3 | 10 | 1 US-08-465-167A-46 | Sequence 46, Appl |
| 11 | 46 | 57.5 | 10 | 1 US-08-465-167A-15 | Sequence 15, Appl |
| 12 | 46 | 57.5 | 10 | 3 US-08-159-339A-595 | Sequence 595, App |
| 13 | 46 | 57.5 | 10 | 4 US-08-627-820-15 | Sequence 15, Appl |
| 14 | 44 | 55.0 | 9 | 1 US-08-465-167A-36 | Sequence 36, Appl |
| 15 | 43 | 53.8 | 9 | 1 US-08-465-167A-31 | Sequence 31, Appl |
| 16 | 43 | 53.8 | 10 | 1 US-08-465-167A-11 | Sequence 11, Appl |
| 17 | 43 | 53.8 | 10 | 1 US-08-465-167A-44 | Sequence 44, Appl |
| 18 | 43 | 53.8 | 10 | 3 US-08-159-339A-592 | Sequence 592, App |
| 19 | 43 | 53.8 | 10 | 4 US-08-627-820-11 | Sequence 11, Appl |
| 20 | 42 | 52.5 | 10 | 1 US-08-465-167A-49 | Sequence 49, Appl |
| 21 | 42 | 52.5 | 330 | 4 US-09-540-236-2570 | Sequence 2570, Ap |
| 22 | 41 | 51.2 | 9 | 1 US-08-465-167A-8 | Sequence 8, Appli |
| 23 | 41 | 51.2 | 9 | 1 US-08-465-167A-27 | Sequence 27, Appl |
| 24 | 41 | 51.2 | 9 | 4 US-08-197-484-93 | Sequence 93, Appl |
| 25 | 41 | 51.2 | 9 | 4 US-08-197-484-152 | Sequence 152, App |
| 26 | 41 | 51.2 | 9 | 4 US-08-627-820-8 | Sequence 8, Appli |
| 27 | 41 | 51.2 | 9 | 4 US-09-543-608A-44 | Sequence 44, Appl |

| | | | | | |
|----|----|------|------|------------------------|-------------------|
| 28 | 41 | 51.2 | 9 | 5 PCT-US95-02121-93 | Sequence 93, Appl |
| 29 | 41 | 51.2 | 9 | 5 PCT-US95-02121-152 | Sequence 152, App |
| 30 | 41 | 51.2 | 157 | 4 US-09-328-352-5371 | Sequence 5371, Ap |
| 31 | 40 | 50.0 | 377 | 4 US-09-328-352-4688 | Sequence 4688, Ap |
| 32 | 39 | 48.8 | 9 | 1 US-08-465-167A-33 | Sequence 33, Appl |
| 33 | 38 | 47.5 | 135 | 4 US-09-489-039A-7753 | Sequence 7753, Ap |
| 34 | 38 | 47.5 | 591 | 4 US-09-540-236-3375 | Sequence 3375, Ap |
| 35 | 37 | 46.2 | 10 | 1 US-08-465-167A-39 | Sequence 39, Appl |
| 36 | 37 | 46.2 | 627 | 4 US-09-403-667A-2 | Sequence 2, Appli |
| 37 | 37 | 46.2 | 627 | 4 US-09-403-667A-4 | Sequence 4, Appli |
| 38 | 36 | 45.0 | 214 | 1 US-08-033-797-3 | Sequence 3, Appli |
| 39 | 36 | 45.0 | 214 | 1 US-08-472-263-3 | Sequence 3, Appli |
| 40 | 36 | 45.0 | 214 | 1 US-08-472-263-3 | Sequence 3, Appli |
| 41 | 36 | 45.0 | 752 | 4 US-09-252-991A-17355 | Sequence 17355, A |
| 42 | 36 | 45.0 | 1064 | 1 US-08-357-598-2 | Sequence 2, Appli |
| 43 | 36 | 45.0 | 1064 | 2 US-09-003-289-2 | Sequence 2, Appli |
| 44 | 36 | 45.0 | 1064 | 5 PCT-US95-16435-2 | Sequence 2, Appli |
| 45 | 36 | 45.0 | 1082 | 1 US-08-357-598-5 | Sequence 5, Appli |

ALIGNMENTS

RESULT 1
US-08-465-167A-1
; Sequence 1, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-1

Query Match 100.0%; Score 80; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
QY      1 EYVIKVSARVRFPPS 16
      |||||||
Db     30 EYVIKVSARVRFPPS 45

RESULT 2
US-08-627-820-1
; Sequence 1, Application US/08627820
; Patent No. 6464980
; GENERAL INFORMATION:
;   APPLICANT: Fikes, John D.
;             Livingston, Brian D.
;             Sette, Alessandro D.
;             Sidney, John C.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
;                   COMPLETE MAGE 1 GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105
; COMPUTER READABLE FORM:
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/627,820
;   FILING DATE: 02-Apr-1996
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/103,623
;   FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Parmelee, Steven W.
;   REGISTRATION NUMBER: 31,990
;   REFERENCE/DOCKET NUMBER: 14137-60
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 467-9600
;   TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 58 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-627-820-1

      Query Match      100.0%; Score 80; DB 4; Length 58;
      Best Local Similarity 100.0%; Pred. No. 2.7e-07;
      Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||||||
Db     30 EYVIKVSARVRFPPS 45

RESULT 3
US-08-465-167A-24
; Sequence 24, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
;   APPLICANT: Fikes, John D.
;   APPLICANT: Livingston, Brian D.
;   APPLICANT: Sette, Alessandro D.
;   APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
```

```
;
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/465,167A
;   FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/103,623
;   FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
;   NAME: Parmelee, Steven W.
;   REGISTRATION NUMBER: 31,990
;   REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 206-467-9600
;   TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 309 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
; US-08-465-167A-24

      Query Match      100.0%; Score 80; DB 1; Length 309;
      Best Local Similarity 100.0%; Pred. No. 1.7e-06;
      Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||||||
Db     281 EYVIKVSARVRFPPS 296

RESULT 4
US-08-993-118-10
; Sequence 10, Application US/08993118
; Patent No. 5997872
; GENERAL INFORMATION:
;   APPLICANT: LUCAS, Sophie;
;   APPLICANT: DE SMET, Charles;
;   APPLICANT: BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
; TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/993,118
;   FILING DATE:
```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
US-08-993-118-10
;
Query Match 100.0%; Score 80; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
   |||||
Db 281 EYVIKVSARVRFPPS 296

RESULT 5
US-08-845-528C-10
; Sequence 10, Application US/08845528C
; Patent No. 6027924
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie;
; APPLICANT: DE SMET, Charles;
; APPLICANT: BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
; TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,528C
; FILING DATE: April 25, 1997
; CLASSIFICATION: 4335
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
US-08-845-528C-10
;
Query Match 100.0%; Score 80; DB 3; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
```

```

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
   |||||
Db 281 EYVIKVSARVRFPPS 296

RESULT 6
US-08-627-820-24
; Sequence 24, Application US/08627820
; Patent No. 6464980
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
; TITLE OF INVENTION: COMPLETE MAGE 1 GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,820
; FILING DATE: 02-Apr-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,623
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
;
US-08-627-820-24
;
Query Match 100.0%; Score 80; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
   |||||
Db 281 EYVIKVSARVRFPPS 296

RESULT 7
US-09-066-281B-10
; Sequence 10, Application US/09066281B
; Patent No. 6475783
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; DE SMET, Charles; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING
; TITLE OF INVENTION: FOR TUMOR REJECTION ANTIGEN PRECURSOR MAGE-C1 AND MAGE-C2
; TITLE OF INVENTION: AND USES THEREOF
```

```
;
;
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,281B
; FILING DATE: April 24, 1998
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455.2 US - JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
US-09-066-281B-10

Query Match 100.0%; Score 80; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
Db 281 EYVIKVSARVRFPPS 296

RESULT 8
US-09-468-433C-10
; Sequence 10, Application US/09468433C
; Patent No. 6680191
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING FOR
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C AN
; TITLE OF INVENTION: MAGE-B FAMILIES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/468,433C
; FILING DATE: December 17, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/066,281
; FILING DATE: April 24, 1998
```

```
;
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5611 JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 662-0200
; TELEFAX: (202) 662-4643
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
US-09-468-433C-10

Query Match 100.0%; Score 80; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
Db 281 EYVIKVSARVRFPPS 296

RESULT 9
US-09-392-714-29
; Sequence 29, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-392-714-29

Query Match 100.0%; Score 80; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
Db 281 EYVIKVSARVRFPPS 296

RESULT 10
US-08-465-167A-46
; Sequence 46, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
```


; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM: Floppy disk
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-46

Query Match 61.3%; Score 49; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVSARVRFFF 14
Db 1 KVSARVRFFF 10

RESULT 11
US-08-465-167A-15
; Sequence 15, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM: Floppy disk
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-15

Query Match 57.5%; Score 46; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVIKVSARVR 11
Db 1 YVIKVSARVR 10

RESULT 12
US-08-159-339A-595
; Sequence 595, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM: Diskette
; MEDIUM TYPE: IBM Compatible
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-005030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 595:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-159-339A-595

Query Match 57.5%; Score 46; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVIKVSARVR 11
| | | | |
Db 1 YVIKVSARVR 10

RESULT 13

US-08-627-820-15
; Sequence 15, Application US/08627820
; Patent No. 6464980

GENERAL INFORMATION:

APPLICANT: Fikes, John D.
Livingston, Brian D.
Sette, Alessandro D.
Sidney, John C.

TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
COMPLETE MAGE 1 GENE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: One Market Plaza, Stuart Street Tower
CITY: San Francisco
STATE: CA

COUNTRY: U.S.A.

ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/627,820

FILING DATE: 02-Apr-1996

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/103,623

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14137-60

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 467-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-08-627-820-15

Query Match 57.5%; Score 46; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVIKVSARVR 11
| | | | |
Db 1 YVIKVSARVR 10

RESULT 14

US-08-465-167A-36
; Sequence 36, Application US/08465167A
; Patent No. 5750395

GENERAL INFORMATION:

APPLICANT: Fikes, John D.
Livingston, Brian D.
Sette, Alessandro D.
Sidney, John C.

TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
IMMUNOGENIC PEPTIDES (as amended)

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 98111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/465,167A

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/103,623

FILING DATE: 06-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14137-60-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-467-9600

TELEFAX: 415-576-0300

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-465-167A-36

Query Match 55.0%; Score 44; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 VSARVRFFF 14
| | | | |
Db 1 VSARVRFFF 9

RESULT 15

US-08-465-167A-31
; Sequence 31, Application US/08465167A
; Patent No. 5750395

GENERAL INFORMATION:

APPLICANT: Fikes, John D.
Livingston, Brian D.
Sette, Alessandro D.
Sidney, John C.

TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
IMMUNOGENIC PEPTIDES (as amended)

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 98111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,167A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/103,623
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-60-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-167A-31

Query Match 53.8%; Score 43; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 KVSARVRFF 13
Db 1 KVSARVRFF 9

Search completed: October 7, 2004, 15:19:25
Job time : 17.4545 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:30:52 ; Search time 59.6364 Seconds
(without alignments)
86.336 Million cell updates/sec

Title: US-09-336-091-4
Perfect score: 80
Sequence: 1 EYVIKVSARVRFPPS 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------------------------|
| 1 | 80 | 100.0 | 30 | 12 | US-10-296-734-1278 Sequence 1278, Ap |
| 2 | 80 | 100.0 | 309 | 9 | US-09-766-889A-2 Sequence 2, Appli |
| 3 | 80 | 100.0 | 309 | 12 | US-10-296-734-828 Sequence 828, App |
| 4 | 80 | 100.0 | 309 | 13 | US-10-085-108-10 Sequence 10, Appl |
| 5 | 80 | 100.0 | 309 | 14 | US-10-177-390-18 Sequence 18, Appl |
| 6 | 80 | 100.0 | 309 | 14 | US-10-160-237-10 Sequence 10, Appl |
| 7 | 80 | 100.0 | 309 | 15 | US-10-117-937-71 Sequence 71, Appl |
| 8 | 80 | 100.0 | 309 | 16 | US-10-657-022-71 Sequence 71, Appl |
| 9 | 80 | 100.0 | 309 | 16 | US-10-741-466-6 Sequence 6, Appli |
| 10 | 80 | 100.0 | 311 | 16 | US-10-741-466-7 Sequence 7, Appli |
| 11 | 80 | 100.0 | 311 | 16 | US-10-741-466-8 Sequence 8, Appli |
| 12 | 80 | 100.0 | 3541 | 12 | US-10-296-734-1454 Sequence 1454, Ap |
| 13 | 64 | 80.0 | 28 | 12 | US-10-296-734-1280 Sequence 1280, Ap |
| 14 | 57 | 71.2 | 12 | 15 | US-10-164-121A-33 Sequence 33, Appl |
| 15 | 57 | 71.2 | 12 | 15 | US-10-164-078A-32 Sequence 32, Appl |

| | | | | | | |
|----|----|------|------|----|----------------------|-------------------|
| 16 | 51 | 63.7 | 317 | 12 | US-10-218-095-2 | Sequence 2, Appli |
| 17 | 51 | 63.7 | 317 | 14 | US-10-157-031-52 | Sequence 52, Appl |
| 18 | 46 | 57.5 | 10 | 15 | US-10-117-937-129 | Sequence 129, App |
| 19 | 46 | 57.5 | 1058 | 16 | US-10-437-963-131272 | Sequence 131272, |
| 20 | 43 | 53.8 | 9 | 16 | US-10-753-158-57 | Sequence 57, Appl |
| 21 | 43 | 53.8 | 12 | 16 | US-10-753-158-56 | Sequence 56, Appl |
| 22 | 41 | 51.2 | 9 | 12 | US-10-149-135-2182 | Sequence 2182, Ap |
| 23 | 41 | 51.2 | 9 | 12 | US-09-935-476-6 | Sequence 6, Appli |
| 24 | 41 | 51.2 | 9 | 14 | US-10-128-711-93 | Sequence 93, Appl |
| 25 | 41 | 51.2 | 9 | 14 | US-10-128-711-152 | Sequence 152, App |
| 26 | 41 | 51.2 | 9 | 14 | US-10-150-797-13 | Sequence 13, Appl |
| 27 | 41 | 51.2 | 9 | 15 | US-10-149-138-4195 | Sequence 4195, Ap |
| 28 | 41 | 51.2 | 9 | 16 | US-10-149-138-4195 | Sequence 4195, Ap |
| 29 | 41 | 51.2 | 592 | 12 | US-10-282-122A-44745 | Sequence 44745, A |
| 30 | 41 | 51.2 | 1070 | 7 | US-08-908-453-11 | Sequence 11, Appl |
| 31 | 41 | 51.2 | 1070 | 12 | US-10-170-385-305 | Sequence 305, App |
| 32 | 40 | 50.0 | 679 | 12 | US-10-282-122A-76599 | Sequence 76599, A |
| 33 | 39 | 48.8 | 9 | 15 | US-10-117-937-128 | Sequence 128, App |
| 34 | 39 | 48.8 | 59 | 16 | US-10-437-963-110118 | Sequence 110118, |
| 35 | 39 | 48.8 | 80 | 14 | US-10-106-698-4953 | Sequence 4953, Ap |
| 36 | 39 | 48.8 | 226 | 12 | US-10-424-599-274271 | Sequence 274271, |
| 37 | 39 | 48.8 | 478 | 9 | US-09-815-242-5719 | Sequence 5719, Ap |
| 38 | 39 | 48.8 | 478 | 9 | US-09-815-242-12480 | Sequence 12480, A |
| 39 | 39 | 48.8 | 478 | 12 | US-10-282-122A-43985 | Sequence 43985, A |
| 40 | 38 | 47.5 | 56 | 11 | US-09-864-408A-3944 | Sequence 3944, Ap |
| 41 | 38 | 47.5 | 56 | 16 | US-10-767-701-47979 | Sequence 47979, A |
| 42 | 38 | 47.5 | 158 | 15 | US-10-074-978A-379 | Sequence 379, App |
| 43 | 38 | 47.5 | 191 | 16 | US-10-767-701-33614 | Sequence 33614, A |
| 44 | 38 | 47.5 | 219 | 15 | US-10-369-493-9785 | Sequence 9785, Ap |
| 45 | 38 | 47.5 | 323 | 16 | US-10-437-963-168711 | Sequence 168711, |

ALIGNMENTS

RESULT 1
US-10-296-734-1278
; Sequence 1278, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1278
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MAGE-1 segment 19
US-10-296-734-1278

Query Match 100.0%; Score 80; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
|||
Db 13 EYVIKVSARVRFPPS 28

RESULT 2
US-09-766-889A-2
; Sequence 2, Application US/09766889A
; Patent No. US20020164654A1
; GENERAL INFORMATION:
; APPLICANT: Luiten, Rosalie

```

; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: van der Bruggen, Pierre
; APPLICANT: Stroobant, Vincent
; APPLICANT: Demotte, Nathalie
; APPLICANT: Schultz, Erwin
; TITLE OF INVENTION: MAGE ANTIGENIC PEPTIDES WHICH BIND HLA-B35 AND HLA-B44
; FILE REFERENCE: L0461/7104
; CURRENT APPLICATION NUMBER: US/09/766,889A
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/177,242
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/243,212
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-766-889A-2

Query Match      100.0%; Score 80; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||||
Db      281 EYVIKVSARVRFPPS 296

RESULT 3
US-10-296-734-828
; Sequence 828, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 828
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MAGE-1 consensus polypeptide
US-10-296-734-828

Query Match      100.0%; Score 80; DB 12; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||||
Db      281 EYVIKVSARVRFPPS 296

RESULT 4
US-10-085-108-10
; Sequence 10, Application US/10085108
; Publication No. US20020176865A1
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING FOR
; TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-Q
; MAGE-B FAMILIES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/085,108
; FILING DATE: 01-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/501,104
; FILING DATE: 09-Feb-2000
; APPLICATION NUMBER: 09/468,433
; FILING DATE: December 17, 1999
; APPLICATION NUMBER: 09/066,281
; FILING DATE: April 24, 1998
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5611.1 JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 318-3400
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-085-108-10

Query Match      100.0%; Score 80; DB 13; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||||
Db      281 EYVIKVSARVRFPPS 296

RESULT 5
US-10-177-390-18
; Sequence 18, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerold
; APPLICANT: N.V. Antwerps Innovatiecentrum
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with Linear
; TITLE OF INVENTION: Polynucleotides by Electroporation
; FILE REFERENCE: 021505wo/JH/ml
; CURRENT APPLICATION NUMBER: US/10/177,390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-18

Query Match      100.0%; Score 80; DB 14; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 281 EYVIKVSARVRFPPS 296

RESULT 6
US-10-160-237-10
; Sequence 10, Application US/10160237
; Publication No. US20030170256A1
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; DE SMET, Charles; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING
; FOR TUMOR REJECTION ANTIGEN PRECURSOR MAGE-C1 AND MAGE-C2
; AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/160,237
; FILING DATE: 04-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,281B
; FILING DATE: April 24, 1998
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455.2 US - JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-160-237-10

Query Match 100.0%; Score 80; DB 14; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 281 EYVIKVSARVRFPPS 296

RESULT 7
US-10-117-937-71
; Sequence 71, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES

; FILE REFERENCE: CTLLMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-117-937-71

Query Match 100.0%; Score 80; DB 15; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 281 EYVIKVSARVRFPPS 296

RESULT 8
US-10-657-022-71
; Sequence 71, Application US/10657022
; Publication No. US20040180354A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Liu, Liping
; APPLICANT: Liu, Zheng
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: MANNK.032A
; CURRENT APPLICATION NUMBER: US/10/657,022
; CURRENT FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: 60/409123
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 610
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-657-022-71

Query Match 100.0%; Score 80; DB 16; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 281 EYVIKVSARVRFPPS 296

RESULT 9
US-10-741-466-6
; Sequence 6, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 6
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Influenza A virus
US-10-741-466-6

Query Match      100.0%; Score 80; DB 16; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||
Db      281 EYVIKVSARVRFPPS 296
      |||

RESULT 10
US-10-741-466-7
; Sequence 7, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 7
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Influenza A virus
US-10-741-466-7

Query Match      100.0%; Score 80; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 7.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||
Db      283 EYVIKVSARVRFPPS 298
      |||

RESULT 11
US-10-741-466-8
; Sequence 8, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 8
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Influenza A virus
US-10-741-466-8

Query Match      100.0%; Score 80; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 7.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||
Db      283 EYVIKVSARVRFPPS 298
      |||
```

```
RESULT 12
US-10-296-734-1454
; Sequence 1454, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1454
; LENGTH: 3541
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Melanoma cancer specific savine
US-10-296-734-1454

Query Match      100.0%; Score 80; DB 12; Length 3541;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFPPS 16
      |||
Db      1704 EYVIKVSARVRFPPS 1719
      |||

RESULT 13
US-10-296-734-1280
; Sequence 1280, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; PRIOR APPLICATION NUMBER: 2003-08-04
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1280
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MAGE-1 segment 20
US-10-296-734-1280

Query Match      80.0%; Score 64; DB 12; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 IKVSARVRFPPS 16
      |||
Db      1 IKVSARVRFPPS 13
      |||

RESULT 14
US-10-164-121A-33
; Sequence 33, Application US/10164121A
; Publication No. US20030228308A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Yi
; APPLICANT: Boon, Thierry
```

; APPLICANT: Van der Bruggen, Pierre
; APPLICANT: Traversari, Catra
; TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-Cw6 Molecules And Uses Thereof
; FILE REFERENCE: LUD-5771
; CURRENT APPLICATION NUMBER: US/10/164,121A
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 33
; LENGTH: 12
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
US-10-164-121A-33

Query Match 71.2%; Score 57; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
|||||
Db 1 EYVIKVSARVRF 12

RESULT 15
US-10-164-078A-32
; Sequence 32, Application US/10164078A
; Publication No. US20030228325A1
; GENERAL INFORMATION:
; APPLICANT: Bilsborough, Janine
; APPLICANT: Schultz, Erwin
; APPLICANT: Panichelli, Christophe
; APPLICANT: Van der Bruggen, Pierre
; APPLICANT: Boon, Thierry
; TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-B18 Molecules And Uses Thereof
; FILE REFERENCE: LUD-5756
; CURRENT APPLICATION NUMBER: US/10/164,078A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 32
; LENGTH: 12
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
US-10-164-078A-32

Query Match 71.2%; Score 57; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
|||||
Db 1 EYVIKVSARVRF 12

Search completed: October 7, 2004, 15:33:47
Job time : 60.6364 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:17:11 ; Search time 18.5455 Seconds
(without alignments)
82.989 Million cell updates/sec

Title: US-09-336-091-4
Perfect score: 80
Sequence: 1 EYVIKVSARVRRFFFP 16
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:**
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 51 | 63.7 | 317 | 2 I38661 | melanoma antigen |
| 2 | 41 | 51.2 | 1070 | 1 A54600 | 1-phosphatidylinos |
| 3 | 40 | 50.0 | 132 | 2 H86416 | unknown protein, 4 |
| 4 | 40 | 50.0 | 327 | 2 F98257 | hypothetical prote |
| 5 | 40 | 50.0 | 327 | 2 AD3027 | hypothetical prote |
| 6 | 40 | 50.0 | 679 | 2 E71282 | probable NH(3)-dep |
| 7 | 39 | 48.8 | 269 | 2 H82322 | probable bis(5'-nu |
| 8 | 39 | 48.8 | 478 | 2 E89790 | 6-phospho-beta-glu |
| 9 | 38 | 47.5 | 392 | 2 S72753 | hypothetical prote |
| 10 | 38 | 47.5 | 402 | 2 H81427 | periplasmic protei |
| 11 | 38 | 47.5 | 402 | 2 S41790 | hypothetical prote |
| 12 | 38 | 47.5 | 610 | 2 T22909 | hypothetical prote |
| 13 | 38 | 47.5 | 1887 | 2 S61703 | fatty-acid synthas |
| 14 | 37.5 | 46.9 | 1801 | 2 T26774 | hypothetical prote |
| 15 | 37 | 46.2 | 152 | 2 H70459 | hypothetical prote |
| 16 | 37 | 46.2 | 296 | 2 AI1195 | hypothetical prote |
| 17 | 37 | 46.2 | 305 | 2 D96769 | ribosomal large ch |
| 18 | 37 | 46.2 | 332 | 2 AI1691 | hypothetical prote |
| 19 | 37 | 46.2 | 332 | 2 AI1319 | oxidoreductases ho |
| 20 | 37 | 46.2 | 357 | 2 H71122 | hypothetical prote |
| 21 | 37 | 46.2 | 430 | 2 T33155 | hypothetical prote |
| 22 | 37 | 46.2 | 534 | 2 S57974 | hypothetical prote |
| 23 | 37 | 46.2 | 769 | 2 H97033 | alpha-glucosidase |
| 24 | 36.5 | 45.6 | 118 | 2 D96788 | protein T4O12.19 { |
| 25 | 36 | 45.0 | 195 | 2 T33188 | hypothetical prote |
| 26 | 36 | 45.0 | 214 | 1 JC4911 | ribosomal protein |
| 27 | 36 | 45.0 | 214 | 2 A42735 | ribosomal protein |
| 28 | 36 | 45.0 | 214 | 2 JC2013 | ribosomal protein |
| 29 | 36 | 45.0 | 247 | 2 T00107 | ABC-type transport |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 30 | 36 | 45.0 | 270 | 2 B86486 | protein F28J9.7 [i |
| 31 | 36 | 45.0 | 288 | 2 T26383 | hypothetical prote |
| 32 | 36 | 45.0 | 301 | 2 AF3504 | acetyl-CoA carboxy |
| 33 | 36 | 45.0 | 336 | 1 G69759 | thioredoxin reduct |
| 34 | 36 | 45.0 | 428 | 2 T21450 | hypothetical prote |
| 35 | 36 | 45.0 | 468 | 2 E69294 | hypothetical prote |
| 36 | 36 | 45.0 | 477 | 2 A70434 | hypothetical prote |
| 37 | 36 | 45.0 | 594 | 2 F72265 | conserved hypothet |
| 38 | 36 | 45.0 | 602 | 2 G81195 | aspartyl-tRNA synt |
| 39 | 36 | 45.0 | 602 | 2 F81831 | aspartate-tRNA lig |
| 40 | 36 | 45.0 | 720 | 2 T02734 | hypothetical prote |
| 41 | 36 | 45.0 | 821 | 2 T40994 | hypothetical prote |
| 42 | 36 | 45.0 | 1124 | 2 A55747 | L-JAK protein-tyro |
| 43 | 36 | 45.0 | 1262 | 2 T13353 | protein stn-B - fr |
| 44 | 36 | 45.0 | 2458 | 2 T17420 | probable polyketid |
| 45 | 35.5 | 44.4 | 122 | 2 G81737 | hypothetical prote |

ALIGNMENTS

RESULT 1
I38661
melanoma antigen MAGE-4 - human
N;Alternate names: MAGE 41 protein; melanoma antigen MAGE-X2
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1996 #sequence revision 07-Jun-1996 #text change 18-Feb-2000
C;Accession: I38661; I38662; PH1297; PH1298; JC2359; G01446
R;De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; B
oon, T.
Immunogenetics 40, 360-369, 1994
A;Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE fa
A;Reference number: I38659; MUID:95012457; PMID:7927540
A;Accession: I38661
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-317 <DEP1>
A;Cross-references: EMBL:U10687; NID:G533514; PIDN:AAA68871.1; PID:G533515
A;Experimental source: antigen MAGE-4a
A;Accession: I38662
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-172,'T',174-317 <DEP2>
A;Cross-references: EMBL:U10688; NID:G533516; PIDN:AAA68872.1; PID:G533517
A;Experimental source: antigen MAGE-4b
R;Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van Pel
J. Exp. Med. 176, 1453-1457, 1992
A;Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolyti
A;Reference number: PH1294; MUID:93018875; PMID:1402688
A;Accession: PH1297
A;Molecule type: DNA
A;Residues: 169-172,'T',174-177 <TRA1>
A;Experimental source: antigen MAGE-4
A;Accession: PH1298
A;Molecule type: DNA
A;Residues: 169-172,'T',174-177 <TRA2>
A;Experimental source: antigen MAGE-41
R;Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
Biochem. Biophys. Res. Commun. 202, 549-555, 1994
A;Title: Cloning and analysis of MAGE-1-related genes.
A;Reference number: JC2358; MUID:94311935; PMID:8037761
A;Accession: JC2359
A;Molecule type: mRNA
A;Residues: 1-172,'T',174-306,'Q',308-317 <DIN>
A;Cross-references: EMBL:U10340; NID:G499123; PIDN:AAA19007.1; PID:G499124
A;Experimental source: melanoma cell line DM150
C;Genetics:
A;Gene: GDB:MAGEA4; MAGE4; MAGE-X2
A;Cross-references: GDB:331119
A;Map position: Xq28-Xq28
A;Introns: #status absent
C;Superfamily: tumor associated protein MAGE
F;169-177/Region: HLA-A1 binding #status predicted

Query Match 63.7%; Score 51; DB 2; Length 317;
Best Local Similarity 56.2%; Pred. No. 0.076;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFFP 16
:|:|:|:|:|:|:|:|:|
Db 289 EHVVRVNARVRIAYPS 304

RESULT 2
A54600
1-phosphatidylinositol 3-kinase (EC 2.7.1.137) 110K chain beta isoform - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A54600
R;Hu, P.; Mondino, A.; Skolnik, E.Y.; Schlessinger, J.
Mol. Cell. Biol. 13, 7677-7688, 1993
A;Title: Cloning of a novel, ubiquitously expressed human phosphatidylinositol 3-kinase
A;Reference number: A54600; MUID:94067128; PMID:8246984
A;Accession: A54600
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1070 <HU1>
A;Cross-references: GB:S67334; NID:g4555759; PIDN:AAB29081.1; PID:g4555760
A;Note: sequence extracted from NCBI backbone (NCBIN:140879, NCBIIP:140880)
C;Genetics:
A;Gene: GDB:PIK3CB; PIK3C1
A;Cross-references: GDB:136233
C;Superfamily: phosphatidylinositol 3-kinase
C;Keywords: phosphotransferase

Query Match 51.2%; Score 41; DB 1; Length 1070;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFF 14
:|:|:|:|:|:|:|:|:|
Db 245 DYVLQVSGRVEYVF 258

RESULT 3
H86416
unknown protein, 44308-43910 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: H86416
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86416
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-132 <STO>
A;Cross-references: GB:AE005172; NID:g10092214; PIDN:AAG12630.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

RESULT 4
F98257
hypothetical protein AGR_L_2017 [imported] - Agrobacterium tumefaciens (strain C58, Cer
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: F98257
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F98257
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89584.1; PID:gl5159473; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_2017
A;Map position: linear chromosome

Query Match 50.0%; Score 40; DB 2; Length 327;
Best Local Similarity 46.7%; Pred. No. 10;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFFP 15
:|:|:|:|:|:|:|:|:|
Db 207 QYFKKMGRLRVRYFMP 221

RESULT 5
AD3027
hypothetical protein Atu3824 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AD3027
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, Y
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AD3027
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44634.1; PID:gl17742257; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3824
A;Map position: linear chromosome

Query Match 50.0%; Score 40; DB 2; Length 327;
Best Local Similarity 46.7%; Pred. No. 10;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFFP 15
:|:|:|:|:|:|:|:|:|
Db 207 QYFKKMGRLRVRYFMP 221

RESULT 6
E71282
probable NH(3)-dependent NAD(+) synthetase (nade) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C;Accession: E71282
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwir
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDe
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998
A;Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: E71282
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-679 <COL>
A;Cross-references: GB:AE001249; GB:AE000520; NID:g3323083; PIDN:AAC65748.1; PID:g332308
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0780

Query Match 50.0%; Score 40; DB 2; Length 679;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 IKVSARVRFFFP 16
: : | | | | : : | | | |
Db 649 VNLSPRVGFYFP 661

RESULT 7
H82322
probable bis(5'-nucleosyl)-tetraphosphatase (symmetrical) (EC 3.6.1.41) VC0441 [similar
C;Species: *Vibrio cholerae*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 31-Dec-2001
C;Accession: H82322
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: H82322
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-269 <HEI>
A;Cross-references: GB:AE004131; GB:AE003852; NID:g9654856; PIDN:AAF93614.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0441
A;Map position: 1
C;Superfamily: bis(5'-nucleosyl)-tetraphosphatase (symmetrical); phosphoesterase core ho
C;Keywords: hydrolase

Query Match 48.8%; Score 39; DB 2; Length 269;
Best Local Similarity 42.9%; Pred. No. 13;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 YVIKVSARVRFFFP 15
: : | | | | : : | | | |
Db 173 YIVNAFTMRFCFP 186

RESULT 8
E89790
6-phospho-beta-glucosidase [imported] - *Staphylococcus aureus* (strain N315)
C;Species: *Staphylococcus aureus*
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: E89790
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguo
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: E89790
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-478 <KUR>
A;Cross-references: GB:BA000018; PID:g13700182; PIDN:BAB41480.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:

A;Gene: bglA
C;Superfamily: Agrobacterium beta-glucosidase

Query Match 48.8%; Score 39; DB 2; Length 478;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 KVSARVRFFFP 15
: : | | | | : : | | | |
Db 257 QIANRLRFFFP 267

RESULT 9
S72753
hypothetical protein B1496 Cl_154 - *Mycobacterium leprae*
C;Species: *Mycobacterium leprae*
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001
C;Accession: S72753; T11012
R;Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A;Description: *Mycobacterium leprae* cosmid B1496.
A;Reference number: S72695
A;Accession: S72753
A;Molecule type: DNA
A;Residues: 1-392 <SMI>
A;Cross-references: EMBL:U00013; NID:g466868; PIDN:AAA17119.1; PID:g466874
R;Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z16918
A;Accession: T11012
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-392 <PAR>
A;Cross-references: EMBL:Z99125; NID:g2398683; PIDN:CAB16170.1; PID:e343547; PID:g23987
C;Genetics:
A;Gene: MLC1536.27c

Query Match 47.5%; Score 38; DB 2; Length 392;
Best Local Similarity 61.5%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 VIKVSARVRFFFP 15
: : | | | | : : | | | |
Db 208 VVRVSANVRFSGP 220

RESULT 10
H81427
periplasmic protein Cj0112 [imported] - *Campylobacter jejuni* (strain NCTC 11168)
C;Species: *Campylobacter jejuni*
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C;Accession: H81427
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilli
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals by
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: H81427
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <PAR>
A;Cross-references: GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB72596.1; PID:g69676
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj0112

Query Match 47.5%; Score 38; DB 2; Length 402;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 VIKVSARVRFFFP 15
: : | | | | : : | | | |
Db 379 VIRVNANKTFYFP 391

RESULT 11
S41790
hypothetical protein 6 - thermophilic bacterium RT8.B4 (fragment)
C:Species: thermophilic bacterium RT8.B4
C:Date: 07-Sep-1994 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
C:Accession: S41790
R:Dwivedi, P.P.; Gibbs, M.D.; Bergquist, P.L.
submitted to the EMBL Data Library, October 1993
A:Description: Cloning, sequencing and over expression of a multifunctional xylanase gene
A:Reference number: S41785
A:Accession: S41790
A:Molecule type: DNA
A:Residues: 1-402 <DWI>
A:Cross-references: EMBL:L18965; NID:g311185; PIDN:AAB42046.1; PID:g552047

Query Match 47.5%; Score 38; DB 2; Length 402;
Best Local Similarity 31.2%; Pred. No. 30;
Matches 5; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRRFFFP 16
| | : | : : : : ||
Db 361 ELVVKITKKINDYPS 376

RESULT 12
T22909
hypothetical protein F58D2.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T22909
R:Sims, M.; Lloyd, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19636
A:Accession: T22909
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-610 <WIL>
A:Cross-references: EMBL:Z81093; PIDN:CAB03147.1; GSPDB:GN00022; CESP:F58D2.2
A:Experimental source: clone F58D2
C:Genetics:
A:Gene: CESP:F58D2.2
A:Map position: 4
A:Introns: 28/1; 93/1; 184/3; 242/3; 422/3; 540/2
C:Superfamily: Caenorhabditis elegans hypothetical protein T08B6.4

Query Match 47.5%; Score 38; DB 2; Length 610;
Best Local Similarity 46.7%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 YVIKVSARVRRFFFP 16
| | : | : : : : ||
Db 395 YTSASSSKLAFFFP 409

RESULT 13
S61703
fatty-acid synthase (EC 2.3.1.85) alpha chain - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein P1409; protein YPL231w
C:Species: Saccharomyces cerevisiae
C:Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 20-Jun-2000
C:Accession: S61703; S65256; S65250; A31107
R:Urrestarazu, L.A.
submitted to the EMBL Data Library, December 1995
A:Reference number: S61699
A:Accession: S61703
A:Molecule type: DNA
A:Residues: 1-1887 <URR>
A:Cross-references: EMBL:X94561; NID:g1181252; PIDN:CAG64256.1; PID:g1181257
R:Urrestarazu, L.A.; Vissers, S.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S65251

A:Accession: S65256
A:Molecule type: DNA
A:Residues: 1-1887 <URW>
A:Cross-references: EMBL:Z73587; NID:g1370477; PIDN:CAA97948.1; PID:g1370478; MIPS:YPL231w
A:Experimental source: strain S288C (AB972)
R:Rieger, M.; Mueller-Auer, S.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S65202
A:Accession: S65250
A:Molecule type: DNA
A:Residues: 1567-1887 <RIE>
A:Cross-references: EMBL:Z73587; MIPS:YPL231w
A:Experimental source: strain S288C (AB972)
R:Mohamed, A.H.; Chirala, S.S.; Mody, N.H.; Huang, W.Y.; Wakil, S.J.
J. Biol. Chem. 263, 12315-12325, 1988
A:Title: Primary structure of the multifunctional alpha subunit protein of yeast fatty acid synthase
A:Reference number: A31107; MUID:88315020; PMID:2900835
A:Accession: A31107
A:Molecule type: DNA
A:Residues: 1-310,'TTGTGG',311-593,'I',595-940,'CLNCVKSWLKLKLERQFP SKLLW','SIRLSMAIALMLH'
A:Cross-references: EMBL:J03936; NID:g171501; PIDN:AAA34601.1; PID:g171502
C:Genetics:
A:Gene: SGD:FAS2
A:Cross-references: SGD:S0006152; MIPS:YPL231w
A:Map position: 16L
C:Superfamily: yeast fatty-acid synthase
C:Keywords: acyltransferase; coenzyme A

Query Match 47.5%; Score 38; DB 2; Length 1887;
Best Local Similarity 58.8%; Pred. No. 1.4e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 EYVIKVSAR----VRFF 13
| | | | | | | : ||
Db 1672 EYVAKVSAREKSAYKFF 1688

RESULT 14
T26774
hypothetical protein Y39G8C.b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26774
R:Smye, R.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20263
A:Accession: T26774
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1801 <WIL>
A:Cross-references: EMBL:AL032634; PIDN:CAB54418.1; GSPDB:GN00019; CESP:Y39G8C.b
A:Experimental source: clone Y39G8C
C:Genetics:
A:Gene: CESP:Y39G8C.b
A:Map position: 1
A:Introns: 69/3; 117/2; 186/3; 226/3; 268/3; 371/3; 540/3; 857/1; 881/3; 1025/3; 1077/1;

Query Match 46.9%; Score 37.5; DB 2; Length 1801;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 3 VIKVSARVRRFF 14
| : | | : | | | | |
Db 840 VLKVSSR-RFF 850

RESULT 15
H70459
hypothetical protein aq_1854 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: H70459
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'Garra, B.;

V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: H70459
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-152 <AQF>
A;Cross-references: GB:AE000759; NID:g2984125; PIDN:AAC07670.1; PID:g2984134; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aq_1854

Query Match 46.2%; Score 37; DB 2; Length 152;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 KVSARVRFFFP 15
:|||||:|:
Db 68 EVSARVKKFYP 78

Search completed: October 7, 2004, 15:18:17
Job time : 20.5455 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:13:16 ; Search time 10.9091 Seconds
(without alignments)
76.370 Million cell updates/sec

Title: US-09-336-091-4
Perfect score: 80
Sequence: 1 EYIKVSARVRRFFFP 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------------------|
| 1 | 80 | 100.0 | 309 | 1 | MAG1_HUMAN P43355 homo sapien |
| 2 | 51 | 63.7 | 317 | 1 | MAG4_HUMAN P43358 homo sapien |
| 3 | 41 | 51.2 | 1070 | 1 | P11B_HUMAN P42338 homo sapien |
| 4 | 41 | 51.2 | 1070 | 1 | P11B_RAT Q92110 rattus norv |
| 5 | 40 | 50.0 | 679 | 1 | NADE_TREPA O83759 treponema p |
| 6 | 39 | 48.8 | 269 | 1 | APAH_VIBCH Q9kus4 vibrio chol |
| 7 | 39 | 48.8 | 1004 | 1 | RPOC_OENOE P95405 oenococcus |
| 8 | 38 | 47.5 | 392 | 1 | YE62_MYCLE Q49682 mycobacteri |
| 9 | 38 | 47.5 | 402 | 1 | TOLB_CAMJE Q9pj14 campylobact |
| 10 | 38 | 47.5 | 402 | 1 | YOR6_CALSR P40983 caldicellul |
| 11 | 38 | 47.5 | 1887 | 1 | FAS2_YEAST P19097 s fatty aci |
| 12 | 37 | 46.2 | 152 | 1 | Y154_AQUAE O67706 aquifex aeo |
| 13 | 37 | 46.2 | 534 | 1 | EK11_YEAST Q03764 saccharomyc |
| 14 | 37 | 46.2 | 591 | 1 | SYD_PSESM Q87y31 pseudomonas |
| 15 | 37 | 46.2 | 632 | 1 | GAAT_HUMAN Q9un88 homo sapien |
| 16 | 36 | 45.0 | 212 | 1 | RL10_PIG Q29195 sus scrofa |
| 17 | 36 | 45.0 | 213 | 1 | RL10_BOVIN Q9xsi3 bos taurus |
| 18 | 36 | 45.0 | 213 | 1 | RL10_HUMAN P27635 homo sapien |
| 19 | 36 | 45.0 | 213 | 1 | RL10_MOUSE P45634 mus musculu |
| 20 | 36 | 45.0 | 269 | 1 | Y163_BUCBP Q89as9 buchnera ap |
| 21 | 36 | 45.0 | 477 | 1 | YF45_AQUAE O67499 aquifex aeo |
| 22 | 36 | 45.0 | 602 | 1 | SYD_NEIMA Q9jt23 neisseria m |
| 23 | 36 | 45.0 | 602 | 1 | SYD_NEIMB Q9k0u5 neisseria m |
| 24 | 36 | 45.0 | 1124 | 1 | JAK3_HUMAN P52333 homo sapien |
| 25 | 36 | 45.0 | 1134 | 1 | A11A_HUMAN P98196 homo sapien |
| 26 | 36 | 45.0 | 1187 | 1 | A11A_MOUSE P98197 mus musculu |
| 27 | 36 | 45.0 | 1262 | 1 | STNB_DROME Q24212 drosophila |
| 28 | 36 | 45.0 | 4074 | 1 | PKHD_HUMAN Q8tcz9 homo sapien |
| 29 | 35.5 | 44.4 | 122 | 1 | Y114_CHLMU Q9pli5 chlamydia m |
| 30 | 35 | 43.8 | 170 | 1 | YF34_AQUAE O67492 aquifex aeo |
| 31 | 35 | 43.8 | 210 | 1 | RL10_CHICK Q08200 gallus gall |
| 32 | 35 | 43.8 | 284 | 1 | FPG_SYNEL P59065 synechococc |
| 33 | 35 | 43.8 | 284 | 1 | FPG_SYNEL P95744 synechococc |

| | | | | |
|------------|---|-----------|------|---------|
| RESULT 1 | | | | |
| MAG1_HUMAN | | | | |
| ID | MAG1_HUMAN | STANDARD; | PRT; | 309 AA. |
| AC | P43355; O00346; | | | |
| DT | 01-NOV-1995 (Rel. 32, Created) | | | |
| DT | 01-NOV-1995 (Rel. 32, Last sequence update) | | | |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update) | | | |
| DE | Melanoma-associated antigen 1 (MAGE-1 antigen) (Antigen M22-E). | | | |
| GN | MAGEA1 OR MAGE1 OR MAGE1A. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=92086861; PubMed=1840703; | | | |
| RA | van der Bruggen P., Traversari C., Chomez P., Lurquin C., de Plaen E. | | | |
| RA | van den Eynde B., Knuth A., Boon T.; | | | |
| RT | "A gene encoding an antigen recognized by cytolytic T lymphocytes on | | | |
| RT | a human melanoma."; | | | |
| RL | Science 254:1643-1647(1991). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Skin; | | | |
| RX | MEDLINE=94311935; PubMed=8037761; | | | |
| RA | Ding M., Beck R.J., Keller C.J., Fenton R.G.; | | | |
| RT | "Cloning and analysis of MAGE-1-related genes."; | | | |
| RL | Biochem. Biophys. Res. Commun. 202:549-555(1994). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=20314869; PubMed=10854409; | | | |
| RA | Mallon A.M., Platzter M., Bates R., Gloeckner G., Botcherby M., | | | |
| RA | Nordsiek G., Strivens M.A., Kioschis P., Dangel A., Cunningham D., | | | |
| RA | Straw R., Weston P., Hunter C., Gilbert M., Fernando S., Goodall K., | | | |
| RA | Kerry G., Greystrom J.S., Clark D., Goerdes M., Blechschmidt K., | | | |
| RA | Rump A., Hinzmann B., Mundy C.R., Miller W., Poustka A., Herman G.E. | | | |
| RA | Rhodes M., Denny P., Rosenthal A., Brown S.D.M.; | | | |
| RT | "Comparative genome sequence analysis of the Bpa/Str region in mouse | | | |
| RT | and man."; | | | |
| RL | Genome Res. 10:758-775(2000). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A., AND VARIANT ALA-32. | | | |
| RA | Chen H., Wang L., Mei M., Qin L., Cong X., Xu J., Wei L., Wang Y., | | | |
| RA | Chen W.; | | | |
| RT | "The polymorphism of MAGE-1 gene in Chinese people."; | | | |
| RL | Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases. | | | |
| RN | [5] | | | |
| RP | MUTAGENESIS. | | | |
| RC | TISSUE=Blood; | | | |
| RX | MEDLINE=94157413; PubMed=8113684; | | | |
| RA | Gaugler B., van den Eynde B., van der Bruggen P., Romero P., | | | |
| RA | Gaforio J.J., de Plaen E., Lethe B., Brasseur F., Boon T.; | | | |
| RT | "Human gene MAGE-3 codes for an antigen recognized on a melanoma by | | | |
| RT | autologous cytolytic T lymphocytes."; | | | |
| RL | J. Exp. Med. 179:921-930(1994). | | | |
| RN | [6] | | | |

Q59603 neisseria g
Q37556 metridium s
P53557 bacillus su
O67748 aquifex aeo
O42586 xenopus lae
O42587 xenopus lae
O51461 borrelia bu
P17980 homo sapien
Q63569 rattus norv
O88685 mus musculu
Q51422 pseudomonas
Q88nj4 pseudomonas

RP SUBCELLULAR LOCATION.
RX MEDLINE=95012905; PubMed=7927954;
RA Schultze-Thater E., Juretic A., Dellabona P., Luscher U., Siegrist W.,
RA Harder F., Heberer M., Zuber M., Spagnoli G.C.;
RT "MAGE-1 gene product is a cytoplasmic protein.";
RL Int. J. Cancer 59:435-439(1994).
CC -!- FUNCTION: Not known, though may play a role in embryonal
CC development and tumor transformation or aspects of tumor
CC progression. Antigen recognized on a melanoma by autologous
CC cytolytic T lymphocytes.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Expressed in many tumors of several types,
CC such as melanoma, head and neck squamous cell carcinoma, lung
CC carcinoma and breast carcinoma, but not in normal tissues except
CC for testes. Never expressed in kidney tumors, leukemias and
CC lymphomas.
CC -!- SIMILARITY: Contains 1 MAGE domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M77481; AAA03229.1; --
DR EMBL; U82670; -; NOT ANNOTATED_CDS.
DR EMBL; AY148486; AAN62752.1; --
DR Genew; HGNC:6796; MAGEA1.
DR MIM; 300016; --
DR GO; GO:0005886; C:plasma membrane; TAS.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Polymorphism; Tumor antigen.
FT DOMAIN 102 301
FT DOMAIN 33 36
FT VARIANT 32 32
FT VARIANT 72 72
FT MUTAGEN 163 163
FT MUTAGEN 169 169
SQ SEQUENCE 309 AA; 34342 MW; 544EEB1F9F4E9D33 CRC64;

Query Match 100.0%; Score 80; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
Db 281 EYVIKVSARVRFPPS 296
|||||

RESULT 2
MAG4_HUMAN
ID MAG4_HUMAN STANDARD; PRT; 317 AA.
AC P43358;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Melanoma-associated antigen 4 (MAGE-4 antigen) (MAGE-x2) (MAGE-41).
GN MAGEA4 OR MAGE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95012457; PubMed=7927540;
RA de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,

RA de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL Immunogenetics 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=94311935; PubMed=8037761;
RA Ding M., Beck R.J., Keller C.J., Fenton R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL Biochem. Biophys. Res. Commun. 202:549-555(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95369706; PubMed=7642112;
RA Imai Y., Shichiho S., Yamada A., Katayama T., Yano H., Itoh K.;
RT "Sequence analysis of the MAGE gene family encoding human tumor-
RT rejection antigens.";
RL Gene 160:287-290(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Duodenum;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: Contains 1 MAGE domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10687; AAA68871.1; --
DR EMBL; U10688; AAA68872.1; --
DR EMBL; U10340; AAA19007.1; --
DR EMBL; D32077; BAA06843.1; --
DR EMBL; BC017723; AAH17723.1; --
DR PIR; I38661; I38661.
DR PDB; 1I4F; 25-JUL-01.
DR Genew; HGNC:6802; MAGEA4.
DR MIM; 300175; --
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Polymorphism; Tumor antigen; 3D-structure.


```

CC      trisphosphate.
CC      -!- PATHWAY: Signaling pathways regulating cell growth.
CC      -!- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC      SUBUNIT (BY SIMILARITY).
CC      -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AJ012482; CAA10046.1; -.
CC      InterPro; IPR008938; ARM.
CC      InterPro; IPR008973; C2_CaLB.
CC      InterPro; IPR000403; PI3_PI4_kinase.
CC      InterPro; IPR002420; PI3K_C2.
CC      InterPro; IPR003113; PI3K_p85B.
CC      InterPro; IPR000341; PI3K_ras_bind.
CC      InterPro; IPR001263; PI3Ka.
CC      Pfam; PF00454; PI3_PI4_kinase; 1.
CC      Pfam; PF00792; PI3K_C2; 1.
CC      Pfam; PF02192; PI3K_p85B; 1.
CC      Pfam; PF00794; PI3K_rbd; 1.
CC      Pfam; PF00613; PI3Ka; 1.
CC      SMART; SM00142; PI3K_C2; 1.
CC      SMART; SM00143; PI3K_p85B; 1.
CC      SMART; SM00144; PI3K_rbd; 1.
CC      SMART; SM00145; PI3Ka; 1.
CC      SMART; SM00146; PI3K; 1.
CC      PROSITE; PS00915; PI3_4_KINASE_1; 1.
CC      PROSITE; PS00916; PI3_4_KINASE_2; 1.
CC      PROSITE; PS00916; PI3_4_KINASE_3; 1.
CC      TRANSFERASE; Kinase; Multigene family.
KW      DOMAIN 800 1050 PI3K/PI4K.
FT      SEQUENCE 1070 AA; 122607 MW; 4E8EB2333E96E4D5 CRC64;
SQ      -----
      Query Match 51.2%; Score 41; DB 1; Length 1070;
      Best Local Similarity 50.0%; Pred. No. 11;
      Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRFFF 14
      :||:|||||:|
Db      245 DYVLQVSGRVEYVF 258

RESULT 5
NADE_TREPA
ID_NADE_TREPA STANDARD; PRT; 679 AA.
AC O83759;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable glutamine-dependent NAD(+) synthetase (EC 6.3.5.1) (NAD(+)
DE synthase [glutamine-hydrolyzing]).
GN NADE OR TP0780.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis

```

```

RT      spirochete.";
RL      Science 281:375-388(1998).
CC      -!- FUNCTION: CAN USE BOTH GLUTAMINE OR AMMONIA AS A NITROGEN
CC      SOURCE (BY SIMILARITY).
CC      -!- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + L-glutamine + H(2)O =
CC      AMP + diphosphate + NAD(+) + L-glutamate.
CC      -!- PATHWAY: NAD biosynthesis.
CC      -!- SIMILARITY: In the C-terminal section; belongs to the NAD
CC      synthetase family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AE001249; AAC65748.1; -.
CC      PIR; E71282; E71282.
CC      TIGR; TP0780; -.
CC      HAMAP; MF_00193; fused; 1.
CC      InterPro; IPR003694; NAD_synthase.
CC      InterPro; IPR003010; Ntlse/CNhydrtse.
CC      Pfam; PF02540; NAD_synthase; 1.
CC      TIGRFAMs; TIGR00552; nade; 1.
KW      Ligase; NAD; ATP-binding; Complete proteome.
FT      DOMAIN 346 679 LIGASE.
FT      NP_BIND 376 383 ATP (BY SIMILARITY).
FT      ACT_SITE 378 378 BY SIMILARITY.
SQ      SEQUENCE 679 AA; 72152 MW; 370BA934CC405878 CRC64;
      Query Match 50.0%; Score 40; DB 1; Length 679;
      Best Local Similarity 53.8%; Pred. No. 10;
      Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      4 IKVSARVRFFFPS 16
      :||:|||||:|
Db      649 VNLSPRVGFYFPS 661

RESULT 6
APAH_VIBCH
ID_APAH_VIBCH STANDARD; PRT; 269 AA.
AC Q9KUS4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Bis(5'-nucleosyl)-tetraphosphatase, symmetrical (EC 3.6.1.41)
DE (Diadenosine tetraphosphatase) (Ap4A hydrolase)
DE P1,P4-tetraphosphate pyrophosphohydrolase).
GN APAH OR VC0441.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E1 Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC      -!- FUNCTION: Hydrolyzes diadenosine 5',5'''-P1,P4-tetraphosphate to
CC      yield ADP (By similarity).
CC      -!- CATALYTIC ACTIVITY: P(1),P(4)-bis(5'-adenosyl)tetraphosphate +

```



```
CC H(2)O = 2 ADP.
CC -1- SIMILARITY: Belongs to the Ap4A hydrolase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE004131; AAF93614.1; -.
CC PIR; H82322; H82322.
CC TIGR; VC0441; -.
CC HAMAP; MF_00199; -.
CC InterPro; IPR004617; ApaH.
CC InterPro; IPR004843; M-ppestrase.
CC InterPro; IPR006186; T_phthase_apaH.
CC Pfam; PF00149; Metallophos; 1.
CC ProDom; PD000252; T_phthase_apaH; 1.
CC TIGRFAMs; TIGR00668; apaH; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 269 AA; 30407 MW; E6D87A1C474D45BB CRC64;

Query Match 48.8%; Score 39; DB 1; Length 269;
Best Local Similarity 42.9%; Pred. No. 6;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 YVIKVSARVFFFP 15
|:::|::|::|
Db 173 YIVNAFTMRMFCFP 186

RESULT 7
RPOC_OENOE STANDARD; PRT; 1004 AA.
AC P95405;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
DE beta' chain) (RNA polymerase beta' subunit) (Fragment).
GN RPOC.
OS Oenococcus oeni (Leuconostoc oenos).
OC Bacteria; Firmicutes; Lactobacillales; Oenococcus.
OX NCBI_TaxID=1247;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCDO 1674;
RX MEDLINE=97016803; PubMed=8863429;
RA Morse R., Collins M.D., O'Hanlon K., Wallbanks S., Richardson P.T.;
RT "Analysis of the beta' subunit of DNA-dependent RNA polymerase does
RT not support the hypothesis inferred from 16S rRNA analysis that
RT Oenococcus oeni (formerly Leuconostoc oenos) is a tachytelic
RT (fast-evolving) bacterium.";
RL Int. J. Syst. Bacteriol. 46:1004-1009(1996).
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: The enzyme consists of the sigma chain and the core
CC enzyme which is composed of 2 alpha chains, 1 beta chain, and 1
CC beta' chain.
CC -1- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; X96384; CAA65248.1; -.
CC HSP; Q9KWU6; iHQM.
CC InterPro; IPR000722; RNA_pol_A.
CC InterPro; IPR007080; RNA_pol_Rpb1_1.
CC InterPro; IPR007066; RNA_pol_Rpb1_3.
CC InterPro; IPR007083; RNA_pol_Rpb1_4.
CC InterPro; IPR007081; RNA_pol_Rpb1_5.
CC InterPro; IPR006592; RNA_polA_N.
CC Pfam; PF04997; RNA_pol_Rpb1_1; 1.
CC Pfam; PF00623; RNA_pol_Rpb1_2; 1.
CC Pfam; PF04983; RNA_pol_Rpb1_3; 1.
CC Pfam; PF05000; RNA_pol_Rpb1_4; 1.
CC Pfam; PF04998; RNA_pol_Rpb1_5; 1.
CC SMART; SM00663; RPOLA_N; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription.
FT NON_TER 1
FT NON_TER 1004 1004
SQ SEQUENCE 1004 AA; 111965 MW; 73750DF47F3A2C36 CRC64;

Query Match 48.8%; Score 39; DB 1; Length 1004;
Best Local Similarity 63.6%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|:::|::|::|
Db 942 EYVLPISARLR 952

RESULT 8
YE62_MYCLE STANDARD; PRT; 392 AA.
AC Q49682;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ML0594.
GN ML0594 OR MLC1536.27C OR U1496A OR B1496_C1_154.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- SIMILARITY: Belongs to the UPF0051 (ycf24) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```


DR EMBL; AL583919; CAC30102.1; --.
DR PIR; S72753; S72753.
DR Leproma; ML0594; --.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF01458; UPF0051; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 392 AA; 42202 MW; 891162F7CA494C6A CRC64;

Query Match 47.5%; Score 38; DB 1; Length 392;
Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 VIKVSARVRFPP 15
|:|||||
Db 208 VVRVSANVRFSGP 220

RESULT 9

TOLB CAMJE
ID TOLB CAMJE STANDARD; PRT; 402 AA.
AC Q9PJ14;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE TolB protein precursor.
GN TOLB OR CJ0112.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
reveals hypervariable sequences."
RL Nature 403:665-668(2000).
CC -!- FUNCTION: Involved in the tonB-independent uptake of proteins (By
similarity).
CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC -!- SIMILARITY: Belongs to the tolB family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AL139074; CAB72596.1; --.
DR PIR; H81427; H81427.
DR HAMAP; MF_00671; -; 1.
DR InterPro; IPR007195; TolB_N.
DR Pfam; PF04052; TolB_N; 1.
KW Transport; Protein transport; Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 17 Potential.
FT CHAIN 18 402 TolB protein.
SQ SEQUENCE 402 AA; 44778 MW; 3E6178CE236B28A6 CRC64;

Query Match 47.5%; Score 38; DB 1; Length 402;
Best Local Similarity 53.8%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 VIKVSARVRFPP 15
|:|||||
Db 379 VIRVNANKTFYFP 391

RESULT 10

YOR6 CALSR
ID YOR6 CALSR STANDARD; PRT; 402 AA.
AC P40983;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in xynA 3'region (ORF6) (Fragment).
OS Caldicellulosiruptor sp. (strain Rt8B.4).
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
OC Caldicellulosiruptor.
OX NCBI_TaxID=28238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97077616; PubMed=8920183;
RA Dwivedi P.P., Gibbs M.D., Saul D.J., Bergquist P.L.;
RT "Cloning, sequencing and overexpression in Escherichia coli of a
xylanase gene, xynA from the thermophilic bacterium Rt8B.4 genus
RT Caldicellulosiruptor."
RL Appl. Microbiol. Biotechnol. 45:86-93(1996).
CC -!- SIMILARITY: SOME, TO M.GENITALIUM MG148.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; L18965; AAB42046.1; --.
DR PIR; S41790; S41790.
KW Hypothetical protein.
FT NON TER 402 402
SQ SEQUENCE 402 AA; 46442 MW; 5D5EC504960641F7 CRC64;

Query Match 47.5%; Score 38; DB 1; Length 402;
Best Local Similarity 31.2%; Pred. No. 14;
Matches 5; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFPPS 16
|:|||||
Db 361 ELVVKITKINDYPS 376

RESULT 11

FAS2 YEAST
ID FAS2 YEAST STANDARD; PRT; 1887 AA.
AC P19097; Q12533;
DT 01-NOV-1990 (Rel. 16, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fatty acid synthase subunit alpha (EC 2.3.1.86) [Includes: Acyl
carrier; 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
(Beta-ketoacyl reductase); 3-oxoacyl-[acyl-carrier protein] synthase
(EC 2.3.1.41) (Beta-ketoacyl synthase)].
GN FAS2 OR YPL231W OR P1409.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88315020; PubMed=2900835;
RA Mohamed A.H., Chirala S.S., Mody N.H., Huang W.Y., Wakil S.J.;
RT "Primary structure of the multifunctional alpha subunit protein of
yeast fatty acid synthase derived from FAS2 gene sequence."
RL J. Biol. Chem. 263:12315-12325(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RA Schueller H.J.;

Query Match 47.5%; Score 38; DB 1; Length 402;
Best Local Similarity 53.8%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 VIKVSARVRFPP 15
|:|||||
Db 379 VIRVNANKTFYFP 391

Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.

[3]

SEQUENCE FROM N.A.

STRAIN=S288C / AB972;

MEDLINE=97313271; PubMed=9169875;

Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoerge W., Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V., Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M., Chung E., Churcher C.M., Coster F., Davis K., Davis R.W., Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A., Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A., Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W., Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K., Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J., Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D., Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D., Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M., Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H., Urrestarazu L.A., Ushinsky S., Vierendeels F., Viissers S., Voss H., Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E., Zhong W.W., Zollner A., Vo D.H., Hani J.;

"The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";

Nature 387:103-105(1997).

[4]

MUTAGENESIS OF GLY-1250.

STRAIN=S288C;

MEDLINE=94316198; PubMed=8041367;

Inokoshi J., Tomoda H., Hashimoto H., Watanabe A., Takeshima H., Omura S.;

"Cerulenin-resistant mutants of Saccharomyces cerevisiae with an altered fatty acid synthase gene.";

Mol. Gen. Genet. 244:90-96(1994).

-!- FUNCTION: Fatty acid synthetase catalyzes the formation of long-chain fatty acids from acetyl-CoA, malonyl-CoA and NADPH.

The alpha subunit contains domains for: acyl carrier protein, 3-oxoacyl-[acyl-carrier protein] reductase, and 3-oxoacyl-[acyl-carrier-protein] synthase. This subunit coordinates the binding of the six beta subunits to the enzyme complex.

-!- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a long-chain acyl-CoA + N CoA + N CO(2) + 2N NADP(+).

-!- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) + [acyl-carrier protein].

-!- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] + NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.

-!- SUBUNIT: [Alpha(6)beta(6)] hexamers of two multifunctional subunits (alpha and beta).

-!- SIMILARITY: TO THE FATTY ACID SYNTHETASE, SUBUNIT ALPHA FROM OTHER FUNGI.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; J03936; AAA34601.1; -.

EMBL; X76890; CAA54218.1; -.

EMBL; X94561; CAA64256.1; -.

EMBL; Z73586; CAA97947.1; -.

EMBL; Z73587; CAA97948.1; -.

PIR; S61703; S61703.

GermOnline; 144213; -.

SGD; S0006152; FAS2.

InterPro; IPR008278; 4-PPT transf.

InterPro; IPR000794; Ketoacyl synth.

InterPro; IPR004568; Pantethn_trn.

InterPro; IPR006162; Ppantne_S.

Pfam; PF01648; ACPS; 1.

Pfam; PF00109; ketoacyl-synt; 1.

Pfam; PF02801; ketoacyl-synt_C; 1.

DR TIGRFAMs; TIGR00556; pantethn trn; 1.

DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.

DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.

KW Fatty acid biosynthesis; Multifunctional enzyme; Oxidoreductase;

KW Transferase; NADP; Phosphopantetheine.

FT DOMAIN 1 ? ACYL CARRIER (ACP).

FT DOMAIN 675 874 BETA-KETOACYL REDUCTASE.

FT DOMAIN 1149 1363 BETA-KETOACYL SYNTHASE.

FT BINDING 180 180 PHOSPHOPANTETHEINE (BY SIMILARITY).

FT ACT_SITE 1305 1305 BETA-KETOACYL SYNTHASE (BY SIMILARITY).

FT MUTAGEN 1250 1250 G->S: CERULENIN-RESISTANCE.

FT CONFLICT 310 310 G -> GTTGTGG (IN REF. 1).

FT CONFLICT 594 594 T -> I (IN REF. 1).

FT CONFLICT 941 1019 AKLRKELVETSEVRKAVSIETALEHKVVNGNSADAAYAQVE

FT IQPRANIQLDFPELPKPKYQVKQIAPAELEGLLDLERVI ->

FT CLNCVKSWLKLKLERQFPKLLWSIRLSMAIALMLHMLKS

FT KFNQELTFNWTSTQNRNHTNRLNKLPLSLRVCWIKELF

FT (IN REF. 1).

FT CONFLICT 1036 1041 RWEMEA -> KMGNGS (IN REF. 1).

FT CONFLICT 1408 1408 A -> S (IN REF. 1).

FT CONFLICT 1671 1671 N -> T (IN REF. 1).

SQ SEQUENCE 1887 AA; 206946 MW; 08B872734CF3AEEA CRC64;

Query Match 47.5%; Score 38; DB 1; Length 1887;

Best Local Similarity 58.8%; Pred. No. 73;

Matches 10; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 EYVIKVSAR----VRFF 13

Db 1672 EYVAKVSAREKSAYKFF 1688

RESULT 12

YI54_AQUAE

ID YI54_AQUAE STANDARD; PRT; 152 AA.

AC O67706;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein AQ_1854 precursor.

GN AQ_1854.

OS Aquifex aeolicus.

OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.

OX NCBI_TaxID=63363;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Graham D.E., Overbeek R., Snead M.A., Keller M., Auja M., Huber R., Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.";

RL Nature 392:353-358(1998).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE000759; AAC07670.1; -.

DR PIR; H70459; H70459.

KW Hypothetical protein; Signal; Complete proteome.

FT SIGNAL 1 16 POTENTIAL.

FT CHAIN 17 152 HYPOTHETICAL PROTEIN AQ_1854.

SQ SEQUENCE 152 AA; 17714 MW; E2ADC83776605AE7 CRC64;

Query Match 46.2%; Score 37; DB 1; Length 152;

Best Local Similarity 63.6%; Pred. No. 8;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 KVSARVRFEEF 15
:|||||:|:|
Db 68 EVSARVKKFYF 78

RESULT 13

ID_EK11 YEAST STANDARD; PRT; 534 AA.
AC Q03764;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ethanolamine kinase (EC 2.7.1.82) (EK).
GN EK11 OR YDR147W OR YD8358.04.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Murphy L., Richards C., Harris D., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [2]

RP CHARACTERIZATION.
RX MEDLINE=99262638; PubMed=10329685;
RA Kim K., Kim K.-H., Storey M.K., Voelker D.R., Carman G.M.;
RA "Isolation and characterization of the Saccharomyces cerevisiae EK11
RT gene encoding ethanolamine kinase.";
RL J. Biol. Chem. 274:14857-14866(1999).
CC -!- FUNCTION: Catalyzes the committed step of phosphatidylethanolamine
CC synthesis via the CDP-ethanolamine pathway. Also exhibits choline
CC kinase activity but its preferred substrate is ethanolamine.
CC -!- CATALYTIC ACTIVITY: ATP + ethanolamine = ADP + O-
CC phosphoethanolamine.
CC -!- PATHWAY: Phosphatidylethanolamine biosynthesis.
CC -!- SIMILARITY: Belongs to the choline/ethanolamine kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; Z50046; CAA90370.1; -.
DR PIR; S57974; S57974.
DR GerMOnline; 140638; -.
DR SGD; S0002554; EK11.
DR GO; GO:0004103; F:choline kinase activity; IDA.
DR GO; GO:0004305; F:ethanolamine kinase activity; IDA.
DR GO; GO:0006646; P:phosphatidylethanolamine biosynthesis; IMP.
DR InterPro; IPR007521; Choline_kin_N.
DR InterPro; IPR002573; Choline_kinase.
DR Pfam; PF04428; Choline_kin_N; 1.
DR Pfam; PF01633; Choline_kinase; 1.
KW Transferase; Kinase.
FT ACT SITE 327 327 BY SIMILARITY.
SQ SEQUENCE 534 AA; 61657 MW; 47B417841D5005A8 CRC64;

Query Match 46.2%; Score 37; DB 1; Length 534;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 13
||:||||:|:
Db 175 EYELKVIARLSFY 187

RESULT 14

SYD_PSESM
ID_SYD_PSESM STANDARD; PRT; 591 AA.
AC Q87Y31;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AsPRS).
GN ASPS OR PSPTO3981.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499;
RA Buell C.R., Jocardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidse T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,
RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AE016870; AAO57440.1; -.
DR TIGR; PSPTO3981; -.
DR HAMAP; MF_00044; -; 1.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_asp.
DR InterPro; IPR004365; tRNA_anti.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 1.
DR Pfam; PF01336; tRNA_anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 591 AA; 66351 MW; 4A841555305F8F36 CRC64;

Query Match 46.2%; Score 37; DB 1; Length 591;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||:||||:
Db 67 EYVVKVVGKVR 77

RESULT 15
ID_GAAT_HUMAN STANDARD; PRT; 632 AA.
AC Q9UN88; Q9NZK8;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Gamma-aminobutyric-acid receptor theta subunit precursor (GABA(A)
DE receptor).
GN GABRQ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 6-632 FROM N.A.
RX MEDLINE=99380615; PubMed=10449790;
RA Bonner T.P., McKernan R.M., Farrar S., le Bourdelles B.,
RA Heavens R.P., Smith D.W., Hewson L., Rigby M.R.,
RA Sirinathsinghji D.J.S., Brown N., Wafford K.A., Whiting P.J.;
RT "Theta, a novel gamma-aminobutyric acid type A receptor subunit.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9891-9896(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20266167; PubMed=10804200;
RA Sinkkonen S.T., Hanna M.C., Kirkness E.F., Korpi E.R.;
RT "GABAA receptor epsilon and theta subunits display unusual structural
RT variation between species and are enriched in the rat locus
RT ceruleus.";
RL J. Neurosci. 20:3588-3595(2000).
CC -!- FUNCTION: GABA, the major inhibitory neurotransmitter in the
CC vertebrate brain, mediates neuronal inhibition by binding to the
CC GABA/benzodiazepine receptor and opening an integral chloride
CC channel.
CC -!- SUBUNIT: Generally pentameric. This subunit coassembles with
CC alpha-2, beta-1 and gamma-1.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF144648; AAD51172.1; -.
DR EMBL; AF189259; AAF70380.1; -.
DR Genew; HGNC:14454; GABRQ.
DR MIM; 300349; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005326; F:neurotransmitter transporter activity; TAS.
DR GO; GO:0004888; F:transmembrane receptor activity; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR008101; GABAAc_receptor.
DR InterPro; IPR006029; Neu_chan_memb.
DR InterPro; IPR006202; Neu_chan_LBD.
DR InterPro; IPR006201; Neu_chan_channel.
DR Pfam; PF02931; Neu_chan_LBD; 1.
DR Pfam; PF02932; Neu_chan_memb; 1.
DR PRINTS; PR01725; GABAARTHETA.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Multigene family; Transmembrane; Chloride channel.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 632 GAMMA-AMINOBUTYRIC-ACID RECEPTOR THETA
FT SUBUNIT.
FT DOMAIN 22 268 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 269 289 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 612 632 POTENTIAL.
FT DISULFID 183 197 BY SIMILARITY.
FT CARBOHYD 127 127 N-LINKED (GLCNAC..) (POTENTIAL).
FT CONFLICT 478 478 F -> I (IN REF. 2).
SQ SEQUENCE 632 AA; 72021 MW; C13267AC9212118B CRC64;

Query Match 46.2%; Score 37; DB 1; Length 632;
Best Local Similarity 53.3%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 EYVIKVSARVRFPP 15
:||| ||| |||
Db 602 DYVPKVDKWSRFLFP 616

Search completed: October 7, 2004, 15:14:00
Job time : 11.9091 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:14:05 ; Search time 57.8182 Seconds
(without alignments)
87.313 Million cell updates/sec

Title: US-09-336-091-4
Perfect score: 80
Sequence: 1 EYVIKVSARVRFPPS 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 80 | 100.0 | 316 | 4 Q8WH6 | Q8wh6 homo sapien |
| 2 | 52 | 65.0 | 458 | 16 Q93N48 | Q93n48 coxiella bu |
| 3 | 51 | 63.7 | 317 | 4 Q14798 | Q14798 homo sapien |
| 4 | 51 | 63.7 | 318 | 4 Q9BUN9 | Q9bun9 homo sapien |
| 5 | 45 | 56.2 | 1126 | 13 Q9PTN6 | Q9ptn6 cyprinus.ca |
| 6 | 43 | 53.8 | 1740 | 5 O97019 | O97019 halocynthia |
| 7 | 41 | 51.2 | 246 | 10 Q8RX05 | Q8rx05 arabidopsis |
| 8 | 41 | 51.2 | 409 | 11 Q8CFF1 | Q8cff1 mus musculu |
| 9 | 41 | 51.2 | 566 | 16 Q8R624 | Q8r624 fusobacteri |
| 10 | 41 | 51.2 | 1064 | 11 Q8BTI9 | Q8bti9 mus musculu |
| 11 | 40 | 50.0 | 132 | 10 Q9C7Q9 | Q9c7q9 arabidopsis |
| 12 | 40 | 50.0 | 201 | 16 Q81S60 | Q81s60 bacillus an |
| 13 | 40 | 50.0 | 248 | 16 Q8KEI5 | Q8kei5 chlorobium |
| 14 | 40 | 50.0 | 264 | 16 Q8F9W3 | Q8f9w3 leptospira |
| 15 | 40 | 50.0 | 327 | 16 Q8U9A4 | Q8u9a4 agrobacteri |
| 16 | 39 | 48.8 | 257 | 2 Q9ALT2 | Q9alt2 campylobact |

| | | | | | |
|----|------|------|------|-----------|--------------------|
| 17 | 39 | 48.8 | 450 | 16 Q88Z80 | Q88z80 lactobacill |
| 18 | 39 | 48.8 | 478 | 16 Q99WV9 | Q99wv9 staphylococ |
| 19 | 39 | 48.8 | 478 | 16 Q8NYG9 | Q8nyg9 staphylococ |
| 20 | 39 | 48.8 | 1106 | 13 O42291 | O42291 gallus gall |
| 21 | 38 | 47.5 | 119 | 16 Q81VY0 | Q81vy0 bacillus an |
| 22 | 38 | 47.5 | 119 | 16 Q81J87 | Q81j87 bacillus ce |
| 23 | 38 | 47.5 | 229 | 5 Q8MPN9 | Q8mpn9 drosophila |
| 24 | 38 | 47.5 | 241 | 5 Q9VTB5 | Q9vtb5 drosophila |
| 25 | 38 | 47.5 | 245 | 5 Q8MPP0 | Q8mpp0 drosophila |
| 26 | 38 | 47.5 | 258 | 2 Q9RMZ5 | Q9rmz5 bacillus an |
| 27 | 38 | 47.5 | 259 | 2 Q8KYF2 | Q8kyf2 bacillus an |
| 28 | 38 | 47.5 | 399 | 16 Q8P1N0 | Q8pln0 streptococc |
| 29 | 38 | 47.5 | 413 | 10 Q9XE32 | Q9xe32 oryza sativ |
| 30 | 38 | 47.5 | 586 | 8 Q9TJR3 | Q9tjr3 prototheca |
| 31 | 38 | 47.5 | 610 | 5 Q9XVM2 | Q9xvm2 caenorhabdi |
| 32 | 37.5 | 46.9 | 504 | 16 Q87MU4 | Q87mu4 vibrio para |
| 33 | 37 | 46.2 | 110 | 4 Q15739 | Q15739 homo sapien |
| 34 | 37 | 46.2 | 119 | 15 Q8Q454 | Q8q454 human immun |
| 35 | 37 | 46.2 | 187 | 8 Q9TCB6 | Q9tcb6 nephroselmi |
| 36 | 37 | 46.2 | 203 | 16 Q81F51 | Q81f51 bacillus ce |
| 37 | 37 | 46.2 | 247 | 10 Q84M44 | Q84m44 oryza sativ |
| 38 | 37 | 46.2 | 268 | 5 Q7YWQ3 | Q7ywg3 caenorhabdi |
| 39 | 37 | 46.2 | 296 | 16 Q8Y8D6 | Q8y8d6 listeria mo |
| 40 | 37 | 46.2 | 305 | 10 Q9C6A2 | Q9c6a2 arabidopsis |
| 41 | 37 | 46.2 | 322 | 10 Q84WG3 | Q84wg3 arabidopsis |
| 42 | 37 | 46.2 | 323 | 12 Q9IGU9 | Q9igu9 tanapox vir |
| 43 | 37 | 46.2 | 323 | 12 Q9DHN9 | Q9dhn9 yaba-like d |
| 44 | 37 | 46.2 | 332 | 16 Q92A47 | Q92a47 listeria in |
| 45 | 37 | 46.2 | 332 | 16 Q8Y5U4 | Q8y5u4 listeria mo |

ALIGNMENTS

RESULT 1

Q8WH6
ID Q8WH6 PRELIMINARY; PRT; 316 AA.
AC Q8WH6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tumor antigen MAGE-N.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hepatoma;
RA Sui Y., Ye J., Wu W.;
RT "Cloning of a new gene of MAGE family in human hepatocellular carcinoma."
RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF443295; AAL37897.1; --
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
SQ SEQUENCE 316 AA; 35409 MW; A463A9A740A089DF CRC64;

Query Match 100.0%; Score 80; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVRFPPS 16
| | | | | | | | | | | | | | | |
Db 288 EYVIKVSARVRFPPS 303

RESULT 2

Q93N48
ID Q93N48 PRELIMINARY; PRT; 458 AA.
AC Q93N48;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DR GO:0005524; F:ATP binding; IEA.
DR GO:004713; F:protein-tyrosine kinase activity; IEA.
DR GO:0016740; F:transferase activity; IEA.
DR GO:0007242; P:intracellular signaling cascade; IEA.
DR GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR000299; Band 4.1.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00069; pkinase; 2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 2.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00295; B41; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00219; TyrcK; 1.
DR PROSITE: PS00057; FERM 3; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 2.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00001; SH2; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1126 AA; 129335 MW; BCF94D35F76C411D CRC64;

Query Match 56.2%; Score 45; DB 13; Length 1126;
Best Local Similarity 76.9%; Pred. No. 14;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 IKVSARVRFFFP 16
DB 95 IKVHYVRFFFS 107

RESULT 6
O97019
ID O97019 PRELIMINARY; PRT; 1740 AA.
AC O97019;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Asc3 protein.
GN ASC3.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyuridae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hepatopancreas;
RX MEDLINE=99101502; PubMed=9886411;
RA Nonaka M., Azumi K., Ji X., Namikawa-Yamada C., Sasaki M., Saiga H.,
Dadds A.W., Sekine H., Homma M.K., Matsushita M., Endo Y., Fujita T.,
RT "Opsonic complement component C3 in the solitary ascidian, Halocynthia
roretzi.";
RL J. Immunol. 162:387-391(1999).
DR EMBL; AB006964; BAA75069.1; -;
DR HSSP; P01024; 1C3D.
DR GO:0005576; C:extracellular; IEA.
DR GO:0003779; F:actin binding; IEA.
DR GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR InterPro: IPR002890; A2M_N.
DR InterPro: IPR001589; ActBind actnin.
DR InterPro: IPR009048; AM_receptor_bind.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR001599; MacrogloblnA2.
DR InterPro: IPR001134; Netrin C.
DR InterPro: IPR008930; Terp_cyc_toroid.
DR Pfam: PF00207; A2M; 1.
DR Pfam: PF01835; A2M_N; 1.
DR Pfam: PF01821; ANATO; 1.
DR Pfam: PF01759; NTR; 1.

DR SMART; SM00643; C345C; 1.
DR PROSITE; PS00019; ACTININ 1; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
SQ SEQUENCE 1740 AA; 196070 MW; 12AB150CDF3B3333 CRC64;

Query Match 53.8%; Score 43; DB 5; Length 1740;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 YVIKVSARVRFFFP 15
DB 375 YIIDMSRSARFFFP 388

RESULT 7
Q8RX05
ID Q8RX05 PRELIMINARY; PRT; 246 AA.
AC Q8RX05;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN AT5G03830.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Full Length cDNA Clones.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY090996; AAM14019.1; -;
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 246 AA; 28058 MW; 0159187F22D4255E CRC64;

Query Match 51.2%; Score 41; DB 10; Length 246;
Best Local Similarity 61.5%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 VIKVSARVRFFFP 15
DB 183 IYKVSVLRLFP 195

RESULT 8
Q8CFF1
ID Q8CFF1 PRELIMINARY; PRT; 409 AA.
AC Q8CFF1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to phosphatidylinositol 3-kinase, catalytic subunit, beta
isoform (Fragment).
DE Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC039650; AAH39650.1; -;

DR GO; GO:0005942; C:1-phosphatidylinositol 3-kinase complex; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016303; F:phosphatidylinositol 3-kinase activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR002420; PI3K_C2.
DR InterPro; IPR003113; PI3K_p85B.
DR InterPro; IPR000341; PI3K_ras_bind.
DR Pfam; PF00792; PI3K_C2; 1.
DR Pfam; PF02192; PI3K_p85B; 1.
DR Pfam; PF00794; PI3K_rbd; 1.
DR SMART; SM00142; PI3K_C2; 1.
DR SMART; SM00143; PI3K_p85B; 1.
DR SMART; SM00144; PI3K_rbd; 1.
KW Kinase.
FT NON_TER 409 409
SQ SEQUENCE 409 AA; 46976 MW; 5545E974376E45D3 CRC64;

Query Match 51.2%; Score 41; DB 11; Length 409;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFFF 14
:|:|:|:|:|:|
Db 239 DYVLQVSGRVEYVF 252

RESULT 9
Q8R624 Q8R624 PRELIMINARY; PRT; 566 AA.
AC Q8R624;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Methyltransferase (EC 2.1.1.-).
GN FN1306.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapattal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018 (2002).
DR EMBL; AE010636; AAL95502.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR003089; AB_hydrolase.
DR InterPro; IPR000051; SAM_bind.
DR InterPro; IPR000379; Ser_estrs.
DR PRINTS; PR00111; ABHYDROLASE.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 566 AA; 66044 MW; 59FBC45E012BF858 CRC64;

Query Match 51.2%; Score 41; DB 16; Length 566;
Best Local Similarity 61.5%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFFF 13
:|:|:|:|:|:|
Db 220 DYVVKNSAEKRFF 232

RESULT 10
Q8BT19 Q8BT19 PRELIMINARY; PRT; 1064 AA.
AC Q8BT19;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphatidylinositol 3-kinase catalytic subunit.
GN PIK3CB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK090116; BAC41102.1; -.
DR MGD; MGI:1922019; Pik3cb.
DR GO; GO:0040016; P:embryonic cleavage; IMP.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR001263; PI3Ka.
DR InterPro; IPR002420; PI3K_C2.
DR InterPro; IPR003113; PI3K_p85B.
DR InterPro; IPR000341; PI3K_ras_bind.
DR InterPro; IPR000403; PI3_Pi4_kinase.
DR Pfam; PF00613; PI3Ka; 1.
DR Pfam; PF00792; PI3K_C2; 1.
DR Pfam; PF02192; PI3K_p85B; 1.
DR Pfam; PF00794; PI3K_rbd; 1.
DR Pfam; PF00454; PI3_Pi4_kinase; 1.
DR SMART; SM00145; PI3Ka; 1.
DR SMART; SM00146; PI3KC; 1.
DR SMART; SM00142; PI3K_C2; 1.
DR SMART; SM00143; PI3K_p85B; 1.
DR SMART; SM00144; PI3K_rbd; 1.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS0290; PI3_4_KINASE_3; 1.
SQ SEQUENCE 1064 AA; 121753 MW; 3E96729B25C52D1C CRC64;

Query Match 51.2%; Score 41; DB 11; Length 1064;
Best Local Similarity 50.0%; Pred. No. 76;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFFF 14
:|:|:|:|:|:|
Db 239 DYVLQVSGRVEYVF 252

RESULT 11
Q9C7Q9 Q9C7Q9 PRELIMINARY; PRT; 132 AA.
AC Q9C7Q9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F15D2.32.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;


```

RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,
RA Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzgier S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC068667; AAG51732.1; -.
DR PIR; H86416; H86416.
DR InterPro; IPR003676; Auxin_inducible.
DR Pfam; PF02519; Auxin_inducible; 1.
KW Hypothetical protein.
SQ SEQUENCE 132 AA; 15103 MW; F516C550FF585A11 CRC64;

Query Match 50.0%; Score 40; DB 10; Length 132;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 YVIKVSARVRFFFP 15
Db :|:|:|:|
36 FVVTSDRIRFAFP 49

RESULT 12
Q81S60 PRELIMINARY; PRT; 201 AA.
AC Q81S60;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN BA1814.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AE017029; AAP25721.1; -.
DR TIGR; BA1814; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 201 AA; 24202 MW; 21B0C8AB1E30FB62 CRC64;

Query Match 50.0%; Score 40; DB 16; Length 201;
Best Local Similarity 38.5%; Pred. No. 24;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 VIKVSARVRFFFP 15
Db :|:|:|:|
51 ILKIASKFRMFFP 63

RESULT 13
Q8KEIS PRELIMINARY; PRT; 248 AA.
ID Q8KEIS
AC Q8KEIS;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc protease, putative.
GN CT0704.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Parksey D.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AE012841; AAM71941.1; -.
DR TIGR; CT0704; -.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002725; DUF45.
DR Pfam; PF01863; DUF45; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 248 AA; 28214 MW; 45382717E1A8A9F9 CRC64;

Query Match 50.0%; Score 40; DB 16; Length 248;
Best Local Similarity 58.3%; Pred. No. 30;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYIKVSARVRF 12
Db :|:|:|:|
14 EYTKVQSQRARY 25

RESULT 14
Q8F9W3 PRELIMINARY; PRT; 264 AA.
ID Q8F9W3
AC Q8F9W3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA0074.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011199; AAN47273.1; -.
KW Hypothetical protein; Complete proteome.

```

SQ SEQUENCE 264 AA; 29314 MW; 818D15A61E4BC22A CRC64;

Query Match 50.0%; Score 40; DB 16; Length 264;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YVIKVSARVRPF 13
| : ||: |||
Db 6 YFYKLEARIRFF 17

RESULT 15

Q8U9A4 PRELIMINARY; PRT; 327 AA.

AC Q8U9A4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Atu3824.
GN ATU3824 OR AGR_L_2017.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Woo L.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Liu F.,
RA Houmieu K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009313; AAL44634.1; --
DR EMBL; AE008301; AAK89584.1; --
DR PIR; AD3027; AD3027.
DR PIR; F98257; F98257.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 327 AA; 38174 MW; FFE53F6F8EAE8B8E3 CRC64;

Query Match 50.0%; Score 40; DB 16; Length 327;
Best Local Similarity 46.7%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRFFFP 15
: | : ||: ||| : |
Db 207 QYFKKMGRLRVYFMP 221

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 12:53:09 ; Search time 47.1818 Seconds
(without alignments)
71.862 Million cell updates/sec

Title: US-09-336-091-7
Perfect score: 57
Sequence: 1 EYVIKVSARVRF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: _geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 57 | 100.0 | 12 | 4 AAB31295 | Aab31295 Peptide d |
| 2 | 57 | 100.0 | 16 | 4 AAB31291 | Aab31291 Peptide d |
| 3 | 57 | 100.0 | 16 | 4 AAB31292 | Aab31292 Peptide d |
| 4 | 57 | 100.0 | 30 | 5 AAU85042 | Aau85042 Human MAG |
| 5 | 57 | 100.0 | 309 | 2 AAR70909 | Aar70909 Human mel |
| 6 | 57 | 100.0 | 309 | 2 AAW81548 | Aaw81548 Tumour re |
| 7 | 57 | 100.0 | 309 | 4 AAB31290 | Aab31290 Amino aci |
| 8 | 57 | 100.0 | 309 | 4 AAEO6806 | Aae06806 Human MAG |
| 9 | 57 | 100.0 | 309 | 5 AAU84814 | Aau84814 Human MAG |
| 10 | 57 | 100.0 | 309 | 6 ABP74195 | Abp74195 Human MAG |
| 11 | 57 | 100.0 | 309 | 6 ABU08930 | Abu08930 Human tum |
| 12 | 57 | 100.0 | 309 | 7 ADC09573 | Adc09573 MAGE-1 pr |
| 13 | 57 | 100.0 | 310 | 6 AAO19742 | Aao19742 Wild-type |
| 14 | 57 | 100.0 | 316 | 6 ABU04419 | Abu04419 Human exp |
| 15 | 57 | 100.0 | 445 | 2 AAY06592 | Aay06592 CLYTA-MAG |
| 16 | 57 | 100.0 | 446 | 2 AAY06590 | Aay06590 Lipoprote |
| 17 | 57 | 100.0 | 1052 | 6 ABR57354 | Abr57354 MatDC16-C |
| 18 | 57 | 100.0 | 3541 | 5 AAU85130 | Aau85130 Human mel |
| 19 | 46 | 80.7 | 10 | 2 AAY38326 | Aay38326 MAGE-deri |
| 20 | 46 | 80.7 | 10 | 2 AAR70948 | Aar70948 Human mel |
| 21 | 46 | 80.7 | 10 | 2 AAR70960 | Aar70960 Human mel |
| 22 | 46 | 80.7 | 10 | 2 AAY46118 | Aay46118 Immunogen |
| 23 | 46 | 80.7 | 10 | 2 AAY45907 | Aay45907 Immunogen |
| 24 | 46 | 80.7 | 10 | 6 ABP74245 | Abp74245 Human MAG |
| 25 | 46 | 80.7 | 10 | 7 ADC09104 | Adc09104 Epitope w |

| | | | | | |
|----|----|------|-----|------------|--------------------|
| 26 | 45 | 78.9 | 10 | 2 AAR70953 | Aar70953 Human mel |
| 27 | 45 | 78.9 | 10 | 2 AAY46179 | Aay46179 Immunogen |
| 28 | 42 | 73.7 | 10 | 2 AAR70957 | Aar70957 Human mel |
| 29 | 42 | 73.7 | 10 | 2 AAR70966 | Aar70966 Human mel |
| 30 | 42 | 73.7 | 10 | 2 AAY46178 | Aay46178 Immunogen |
| 31 | 41 | 71.9 | 9 | 2 AAR73839 | Aar73839 Antigen f |
| 32 | 41 | 71.9 | 9 | 2 AAR78916 | Aar78916 MAGE 1 28 |
| 33 | 41 | 71.9 | 9 | 2 AAR70923 | Aar70923 Human mel |
| 34 | 41 | 71.9 | 9 | 2 AAY47412 | Aay47412 Immunogen |
| 35 | 41 | 71.9 | 9 | 4 AAB98466 | Aab98466 HLA class |
| 36 | 41 | 71.9 | 9 | 4 AAG84522 | Aag84522 Human leu |
| 37 | 41 | 71.9 | 9 | 4 AAU06282 | Aau06282 Human leu |
| 38 | 41 | 71.9 | 9 | 4 AAG88274 | Aag88274 Human leu |
| 39 | 41 | 71.9 | 9 | 4 AAG89370 | Aag89370 Human leu |
| 40 | 41 | 71.9 | 9 | 4 AAJ00058 | Aaj00058 Hepatitis |
| 41 | 41 | 71.9 | 9 | 5 ABB78054 | Abb78054 Amino aci |
| 42 | 41 | 71.9 | 9 | 5 AAE31181 | Aae31181 Human mag |
| 43 | 41 | 71.9 | 9 | 6 ABB99586 | Abb99586 Peptide M |
| 44 | 41 | 71.9 | 28 | 5 AAU85043 | Aau85043 Human MAG |
| 45 | 41 | 71.9 | 157 | 6 ADA34084 | Ada34084 Acinetoba |

ALIGNMENTS

RESULT 1
AAB31295
ID AAB31295 standard; peptide; 12 AA.
XX
AC AAB31295;
XX
DT 20-APR-2001 (first entry)
XX
DE Peptide derived from human MAGE-A1 HLA class II-binding protein.
XX
KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;
KW MAGE-A1 HLA class II-binding protein; vaccine.
XX
OS Homo sapiens.
XX
PN WO200078806-A1.
XX
PD 28-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US016287.
XX
PR 18-JUN-1999; 99US-00336091.
XX
(LUDW-) LUDWIG INST CANCER RES.
Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;
WPI; 2001-102698/11.
Novel MAGE-A1 human leukocyte antigen class II peptides which bind to and are presented to the class II molecules, useful for inducing immune response and treating cancers characterized by expression of MAGE-A1.
Claim 3; Page 43; 78pp; English.
The present sequence is derived from a human MAGE-A1 HLA (human leukocyte antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA binding protein stimulate the activity and proliferation of CD4+ T lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic agent for diagnosing a disorder characterized by expression of MAGE-A1. The protein is used for treating a disorder characterized by expression of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas, colorectal carcinomas, osteosarcomas, and lymphocytic leukemias. Peptides derived from the MAGE-A1 HLA binding protein are useful in the production of anti-tumour vaccines
Sequence 12 AA;

PR 26-MAY-2000; 2000AU-00007761.
XX (AUSU) UNIV AUSTRALIAN NAT.
PA Thomson SA, Ramshaw IA;
XX WPI; 2002-147575/19.
DR N-PSDB; ABK36862.
XX
PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.
XX
PS Example 3; Fig 27; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
CC a peptide derived from a parent protein used to construct a savine of the
XX invention
SQ Sequence 30 AA;

Query Match 100.0%; Score 57; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
Db 13 EYVIKVSARVRF 24
RESULT 5
AAR70909
ID AAR70909 standard; protein; 309 AA.
XX
AC AAR70909;
XX
DT 25-MAR-2003 (revised)
DT 09-OCT-1995 (first entry)
XX
DE Human melanoma antigen MAGE-1.
XX
KW Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW HLA-restricted cytotoxic T-lymphocyte activity.
XX
OS Homo sapiens.
XX
PN WO9504542-A1.
XX
PD 16-FEB-1995.
XX
PF 02-AUG-1994; 94WO-US008721.
XX

PR 06-AUG-1993; 93US-00103623.
XX (CYTE-) CYTEL CORP.
PA Fikes JD, Livingston BD, Sette AD, Sidney JC;
PI WPI; 1995-090681/12.
XX N-PSDB; AAQ85435.
DR
XX Human melanoma antigen, MAGE-1, peptide(s) - useful for stimulating
PT immune response against melanoma.
PT
XX Example 1; Fig 1; 59pp; English.
PS
XX AAQ85435 encodes AAR70909 human melanoma antigen MAGE-1, it was used to
CC produce the C-terminal MAGE-1 peptides described in AAR70915 to AAR70969.
CC These peptides are useful for defining epitopes that engender a HLA-
CC restricted cytotoxic lymphocyte activity against MAGE-1 antigens.
CC Compsns. containing these peptides can be administered, as a vaccine to
CC patients susceptible to MAGE associated tumours, e.g. melanomas. (Updated
CC on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 309 AA;
Query Match 100.0%; Score 57; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EYVIKVSARVRF 12
Db 281 EYVIKVSARVRF 292
RESULT 6
AAW81548
ID AAW81548 standard; protein; 309 AA.
XX
AC AAW81548;
XX
DT 01-MAR-1999 (first entry)
XX
DE Tumour rejection antigen precursor MAGE-A1.
XX
KW MAGE-A1; human; tumour rejection antigen precursor; TRAP; therapy;
KW diagnosis.
XX
OS Homo sapiens.
XX
PN WO9849184-A1.
XX
PD 05-NOV-1998.
XX
PF 24-APR-1998; 98WO-US008493.
XX
PR 25-APR-1997; 97US-00845528.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Lucas S, De Smet C, Boon-Falleur T;
XX
DR WPI; 1999-024041/02.
DR N-PSDB; AAV69719.
XX
PT Tumour rejection antigen precursors - used for determining presence of
PT cytolytic T cells specific for complexes of a human leukocyte antigen.
XX
PS Disclosure; Page 50-51; 84pp; English.
XX
CC This is the amino acid sequence of human tumour rejection antigen
CC precursor (TRAP) MAGE-A1. MAGE-A1 cDNA (see AAV69719) shows homology to
CC novel human MAGE-C1 cDNA (see AAV69720). MAGE-C1 (see AAW81546) is a
CC novel member of the MAGE family that may be recognised by cytotoxic T
CC cells, leading to lysis of the tumour cells which express it. It is

CC expressed in a variety of tumours and in normal testis cells, but not by
CC other normal cells. The invention provides MACE-C1 and MACE-C2 nucleic
CC acids and polypeptides, useful e.g. in a claimed method for determining
CC the presence of cytolytic T cells specific for complexes of a human
CC leukocyte antigen (HLA)
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 57; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
|||
Db 281 EYVIKVSARVRF 292

RESULT 7
AAB31290
ID AAB31290 standard; protein; 309 AA.
XX
AC AAB31290;
XX
DT 20-APR-2001 (first entry)
XX
DE Amino acid sequence of human MAGE-A1 HLA class II-binding protein.
XX
KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;
KW MAGE-A1 HLA class II-binding protein; vaccine.
XX
OS Homo sapiens.
XX
PN WO200078806-A1.
XX
PD 28-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US016287.
XX
PR 18-JUN-1999; 99US-00336091.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX

PI Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;
XX
DR WPI; 2001-102698/11.
DR N-PSDB; AAF24676.
XX
PT Novel MAGE-A1 human leukocyte antigen class II peptides which bind to and
PT are presented to the class II molecules, useful for inducing immune
PT response and treating cancers characterized by expression of MAGE-A1.
XX
PS Claim 1; Page 63; 78pp; English.
XX
CC The present sequence represents a human MAGE-A1 HLA (human leukocyte
CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA
CC binding protein stimulate the activity and proliferation of CD4+ T
CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic
CC agent for diagnosing a disorder characterized by expression of MAGE-A1.
CC The protein is used for treating a disorder characterized by expression
CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,
CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias. Peptides
CC derived from the MAGE-A1 HLA binding protein are useful in the production
CC of anti-tumour vaccines
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 57; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
|||
Db 281 EYVIKVSARVRF 292

RESULT 8
AAE06806
ID AAE06806 standard; protein; 309 AA.
XX
AC AAE06806;
XX

DT 16-OCT-2001 (first entry)
XX
DE Human MAGE-A1 protein.
XX

KW MAGE antigenic peptide; Human leukocyte antigen; HLA-B35; HLA-B44;
KW tumour cell; immunostimulant; antigen presentation; cancer; melanoma;
KW CD8+ cytotoxic T lymphocyte; colorectal; prostate; gastric carcinoma;
KW myeloma; brain tumour; sarcoma; seminoma; ovarian tumour; cytostatic;
KW gene therapy; human; MAGE-A1; tumour rejection antigen; TRA.
XX

OS Homo sapiens.
XX
PN WO200153833-A1.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US002008.
XX
PR 20-JAN-2000; 2000US-0177242P.
PR 25-OCT-2000; 2000US-0243212P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX

PI Luiten R, Boon-Falleur T, Van Der Bruggen P, Stroobant V;
PI Demotte N, Schultz E;
XX
DR WPI; 2001-488724/53.
DR N-PSDB; AAD12987.
XX

PT Functional variants and isolated mimetics of a MAGE-A1 HLA-B35 or HLA-B44
PT binding peptide, or of a MAGE-A3 HLA-B35 binding peptide, used in
PT diagnosis and treatment of a disorder characterized by expression of MAGE
PT -A1 or -A3.
XX

PS Claim 2; Page 86-87; 103pp; English.

XX
CC The invention relates to functional variants and isolated mimetics of a
CC MAGE-A1 human leukocyte antigen (HLA)-B35 or HLA-B44 binding peptide, or
CC of a MAGE-A3 HLA-B35 binding peptide, identified by methods described in
CC the specification. MAGE genes encode tumour rejection antigens (TRAs)
CC presented to T lymphocytes by HLA-B35 and HLA-B44 molecules. The MAGE
CC antigenic peptide acts by binding to HLA molecules on tumour cells and
CC stimulating recognition of these cells and thus signalling them to the
CC immune system for destruction. The peptide when presented by HLA molecule
CC induces the activation and stimulation of CD8+ cytotoxic T lymphocytes.
CC The MAGE antigenic peptide is used to treat and diagnose disorders
CC characterised by expression of MAGE-A1 or -A3. Disorders include cancers
CC e.g melanomas, oesophageal, lung, head and neck, breast, colorectal,
CC prostate, renal, bladder, hepatocellular, papillary thyroid and gastric
CC carcinomas, myelomas, brain tumours, sarcomas, seminomas, and ovarian
CC tumours. The present sequence is human MAGE-A1 protein
XX

SQ Sequence 309 AA;

Query Match 100.0%; Score 57; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
|||
Db 281 EYVIKVSARVRF 292

RESULT 9
AAU84814

ID AAU84814 standard; protein; 309 AA.
XX
AC AAU84814;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human MAGE-1 consensus sequence.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Homo sapiens.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU0000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
XX
PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.
XX
PS Example 3; Fig 27; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
CC a consensus sequence for a parent protein used to design a savine of the
XX invention
SQ Sequence 309 AA;

Query Match 100.0%; Score 57; DB 5; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
| | | | | | | | | |
Db 281 EYVIKVSARVRF 292

RESULT 10
ABP74195

ID ABP74195 standard; protein; 309 AA.
XX
AC ABP74195;
XX
DT 03-FEB-2003 (first entry)
XX
DE Human MAGE-1 protein SEQ ID.NO:71.
XX
KW Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;
KW T cell.
XX
OS Homo sapiens.
XX
PN WO200281646-A2.
XX
PD 17-OCT-2002.
XX
PF 04-APR-2002; 2002WO-US011101.
XX
PR 06-APR-2001; 2001US-0282211P.
PR 07-NOV-2001; 2001US-0337017P.
PR 07-MAR-2002; 2002US-0363210P.
XX
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX
PI Simard JLL, Diamond DC, Liu L, Xie Z;
XX
DR WPI; 2003-067518/06.
DR N-PSDB; ABQ83847.
XX
PT Novel epitopes useful as vaccines, comprises peptides or nucleic acid
PT encoding the peptides, that are useful epitopes of target-associated
PT antigens.
XX
PS Claim 1; Page 156; 352pp; English.
XX
CC The present invention describes an isolated epitope (I) and an epitope
CC cluster. Also described is a vaccine or immunotherapeutic composition
CC (VC) comprising (I). (I) has cytostatic activity. VC is useful for
CC treating an animal, by administering to an animal the vaccine or
CC immunotherapeutic composition. VC is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic composition, by
CC administering VC to an HLA-transgenic animal and evaluating
CC immunogenicity based on a characteristic of the animal, or by in vitro
CC primary stimulation of a T cell and evaluating immunogenicity. (I) is
CC useful for determining specific T cell frequency, by contacting T cells
CC with a MHC-peptide complex, and further comprises ELISPOT analysis,
CC limiting dilution analysis, flow cytometry, in situ hybridisation and/or
CC polymerase chain reaction (PCR). ABQ83843 to ABQ83858 and ABP74128 to
CC ABP74713 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 57; DB 6; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
| | | | | | | | | |
Db 281 EYVIKVSARVRF 292

RESULT 11
ABU08930
ID ABU08930 standard; protein; 309 AA.
XX
AC ABU08930;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human tumour rejection antigen precursor, MAGE-A1.
XX

KW TRAP; tumour rejection antigen precursor; cytolytic T-cell; CTL; tumour;
KW seminoma; bladder transitional-cell carcinoma; NSCLC; adaptor;
KW head-and-neck squamous-cell carcinoma; breast carcinoma; sarcoma;
KW cutaneous melanoma; nonsmall cell lung cancer; MAGE-A1; human.
XX
OS Homo sapiens.
XX
PN US2002176865-A1.
XX PD
XX PD 28-NOV-2002.
XX
PF 01-MAR-2002; 2002US-00085108.
XX
PR 25-APR-1997; 97US-00845528.
PR 24-APR-1998; 98US-00066281.
PR 17-DEC-1999; 99US-00468433.
PR 09-FEB-2000; 2000US-00501104.
XX
PA (LUCA/) LUCAS S.
PA (BOON/) BOON-FALLEUR T.
XX
PI Lucas S, Boon-Falleur T;
XX
DR WPI; 2003-328468/31.
DR N-PSDB; ABX93696.
XX
PT Novel isolated nucleic acid encoding tumor rejection antigen precursor
PT MAGE-C3, MAGE-B5, or MAGE-B6, useful as diagnostic probes to determine
PT presence of abnormal e.g., tumor cells expressing MAGE-C1, MAGE-B5 or
PT MAGE-B6.
XX
PS Disclosure; Fig 2; 59pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule which encodes
CC a tumour rejection antigen precursor (TRAP) having an amino acid sequence
CC of a TRAP encoded by a fully defined MAGE-C3, MAGE-B5, or MAGE-B6
CC polynucleotide sequence. Also disclosed is a method which is useful for
CC determining presence of cytolytic T-cells specific for complexes of human
CC leukocyte antigen (HLA) and a peptide derived from the nucleic acid in a
CC cytotoxic T-lymphocyte (CTL)-containing sample. The nucleic acid is
CC useful as a diagnostic probe to determine the presence of abnormal
CC (tumour) cells such as seminoma, bladder transitional-cell carcinoma,
CC head-and-neck squamous-cell carcinoma, breast carcinoma, sarcoma,
CC cutaneous melanoma or nonsmall cell lung cancer (NSCLC) which express
CC MAGE-C1, MAGE-B5 or MAGE-B6. The nucleic acid is useful for diagnosing a
CC disorder characterised by expression of MAGE-C1, MAGE-B5 or MAGE-B6 TRAPs
CC or tumour rejection antigens (TRAS). The present sequence represents the
CC amino acid sequence of the human tumour rejection antigen precursor, MAGE
CC -A1
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 57; DB 6; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
Db 281 EYVIKVSARVRF 292

RESULT 12
ADC09573
ID ADC09573 standard; protein; 309 AA.
XX
AC ADC09573;
XX
DT 18-DEC-2003 (first entry)
XX
DE MAGE-1 protein #SEQ ID 71.
XX
KW Epitope; immunological; vaccine;
KW major histocompatibility complex class I; MHC class I; cancer;

KW immunisation.
XX Unidentified.
OS
XX WO2003008537-A2.
XX
PD 30-JAN-2003.
XX
PF 29-MAR-2002; 2002WO-US010189.
XX
PR 06-APR-2001; 2001US-0282211P.
PR 07-NOV-2001; 2001US-0337017P.
PR 07-MAR-2002; 2002US-0363210P.
XX
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX
PI Simard J JL, Diamond DC, Liu L, Xie Z;
XX
DR WPI; 2003-248010/24.
XX
PT Epitope having high affinity for major histocompatibility complex class I
PT useful for treating an animal, evaluating immunogenicity of a vaccine or
PT therapeutic composition and for diagnosing a disease.
XX
PS Claim 1; SEQ ID NO 71; 239pp; English.
XX
CC The invention relates to an isolated epitope polypeptide that has high
CC affinity for major histocompatibility complex (MHC) class I, and an
CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine
CC or immunotherapeutic composition containing an epitope of the invention.
CC Compositions of the invention may be used in the treatment of cancer. The
CC method can be combined with a radiation therapy, chemotherapy,
CC biochemotherapy or surgery. The composition is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
CC -peptide complexes of the invention are useful for determining specific T
CC cell frequency. This method is useful for evaluating immunological
CC response, by performing the method prior to and subsequent to an
CC immunisation step. Compositions of the invention are useful for
CC diagnosing a disease. The current sequence represents an epitope of the
CC invention with high affinity for MHC class I.
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 57; DB 7; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
Db 281 EYVIKVSARVRF 292

RESULT 13
AAO19742
ID AAO19742 standard; protein; 310 AA.
XX
AC AAO19742;
XX
DT 11-AUG-2003 (first entry)
XX
DE Wild-type MAGE1 protein.
XX
KW Stabilised mRNA; translation optimised; vaccine; tissue repair;
KW sequence modification determination; gene therapy; cytostatic; virucide;
KW antibacterial; protozoacide; nootropic; neuroprotective; infection;
KW antiparkinsonian; immunostimulant; cancer; MAGE1 protein.
XX
OS Unidentified.
XX
PN WO200298443-A2.
XX
PD 12-DEC-2002.

PF 05-JUN-2002; 2002WO-EP006180.
 XX
 PR 05-JUN-2001; 2001DE-01027283.
 XX
 PA (VMUE/) VON DER MUELBE F.
 XX
 PI Von Der Muelbe F, Hoerr I, Pascolo S;
 XX
 DR WPI; 2003-148621/14.
 DR N-PSDB; ABZ69107.
 XX
 PT Composition containing mRNA modified for optimal translation and
 PT stability, useful for treating e.g. tumors or infections, comprises
 PT increased G/C content and fewer rare codons.
 XX
 PS Disclosure; Fig 2B; 75pp; German.
 XX
 CC The present invention relates to a pharmaceutical composition containing
 CC at least one modified RNA encoding a biologically active or antigenic
 CC protein. The RNA is modified to optimise translation of the sequence. The
 CC compositions are used for vaccination against a wide range of infectious
 CC diseases (viral, bacterial or protozoal) or cancer, or for tissue
 CC regeneration, e.g. in cases of Alzheimer's or Parkinson's diseases and
 CC arthritis, but also to express proteins such as dystrophins, chloride ion
 CC channels (for treating cystic fibrosis) and enzymes (either for treating
 CC metabolic disorders or for synthesis of neurotransmitters such as
 CC dopamine). The present sequence is the wild-type MAGE1 protein
 XX
 SQ Sequence 310 AA;

Query Match 100.0%; Score 57; DB 6; Length 310;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
 |||||
 Db 280 EYVIKVSARVRF 291

RESULT 14
 ABU04419
 ID ABU04419 standard; protein; 316 AA.
 XX
 AC ABU04419;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE Human expressed protein tag (EPT) #1085.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX

OS Homo sapiens.
 XX
 PN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Chiciz RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.
 DR
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 PS Example 2; SEQ ID NO 1085; 134pp; English.
 XX
 CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, or protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 316 AA;

Query Match 100.0%; Score 57; DB 6; Length 316;
 Best Local Similarity 100.0%; Pred. No. 0.004;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
 |||||
 Db 288 EYVIKVSARVRF 299

RESULT 15
 AAY06592
 ID AAY06592 standard; protein; 445 AA.
 XX
 AC AAY06592;
 XX
 DT 26-OCT-1999 (first entry)
 XX
 DE CLYTA-MAGE-1-His fusion protein.

XX MAGE-1; CLYTA-MAGE-1-His; fusion protein; tumour; melanoma;
 KW breast cancer; bladder cancer; lung cancer; colon cancer;
 KW head and squamous cell carcinoma; oesophagus carcinoma; vaccine; human.
 XX

OS Streptococcus pneumoniae.
 OS Homo sapiens.
 OS Synthetic.
 OS Chimeric.
 XX
 PN WO9940188-A2.
 XX
 PD 12-AUG-1999.
 XX
 PF 02-FEB-1999; 99WO-EP000660.
 XX
 PR 05-FEB-1998; 98GB-00002543.
 PR 06-FEB-1998; 98GB-00002650.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Cabezon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;

XX WPI; 1999-494293/41.
 DR N-PSDB; AAX87593.

XX New protein derivatives used in cancer vaccine therapy for treating a
PT range of cancers including melanomas, carcinomas and cancers of breast.
XX
PS Example 9; Page 69-70; 72pp; English.
XX
CC The present sequence represents a fusion protein composed of the C-
CC terminal portion of the Streptococcus pneumoniae LYTA protein (CLYTA),
CC the human MAGE-1 tumour-associated antigen and a hexahistidine tail. A
CC vector designed for recombinant expression of the fusion protein in
CC Escherichia coli is provided. The CLYTA moiety provides expression of
CC soluble fusion protein, facilitates affinity purification, and also acts
CC as a T-helper epitope. The invention relates to MAGE proteins fused to an
CC immunological fusion partner, e.g. CLYTA-MAGE-1-His. These novel fusion
CC proteins provide vaccines for immunotherapy of melanomas or other MAGE-
CC associated tumours like breast, bladder, lung and non-small cell lung
CC cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus
CC carcinoma
XX
SQ Sequence 445 AA;

Query Match 100.0%; Score 57; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
| | | | | | | | | |
Db 408 EYVIKVSARVRF 419

Search completed: October 7, 2004, 15:13:03
Job time : 48.1818 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:18:26 ; Search time 13.0909 Seconds
(without alignments)
47.324 Million cell updates/sec

Title: US-09-336-091-7
Perfect score: 57
Sequence: 1 EVYIKVSARVF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 57 | 100.0 | 58 | 1 | US-08-465-167A-1 |
| 2 | 57 | 100.0 | 58 | 4 | US-08-627-820-1 |
| 3 | 57 | 100.0 | 309 | 1 | US-08-465-167A-24 |
| 4 | 57 | 100.0 | 309 | 2 | US-08-993-118-10 |
| 5 | 57 | 100.0 | 309 | 3 | US-08-845-528C-10 |
| 6 | 57 | 100.0 | 309 | 4 | US-08-627-820-24 |
| 7 | 57 | 100.0 | 309 | 4 | US-09-066-281B-10 |
| 8 | 57 | 100.0 | 309 | 4 | US-09-468-433C-10 |
| 9 | 57 | 100.0 | 309 | 4 | US-09-392-714-29 |
| 10 | 46 | 80.7 | 10 | 1 | US-08-465-167A-15 |
| 11 | 46 | 80.7 | 10 | 3 | US-08-159-339A-595 |
| 12 | 46 | 80.7 | 10 | 4 | US-08-627-820-15 |
| 13 | 42 | 73.7 | 10 | 1 | US-08-465-167A-49 |
| 14 | 41 | 71.9 | 9 | 1 | US-08-465-167A-8 |
| 15 | 41 | 71.9 | 9 | 1 | US-08-465-167A-27 |
| 16 | 41 | 71.9 | 9 | 4 | US-08-197-484-93 |
| 17 | 41 | 71.9 | 9 | 4 | US-08-197-484-152 |
| 18 | 41 | 71.9 | 9 | 4 | US-08-627-820-8 |
| 19 | 41 | 71.9 | 9 | 4 | US-09-543-608A-44 |
| 20 | 41 | 71.9 | 9 | 5 | PCT-US95-02121-93 |
| 21 | 41 | 71.9 | 9 | 5 | PCT-US95-02121-152 |
| 22 | 41 | 71.9 | 157 | 4 | US-09-328-352-5371 |
| 23 | 39 | 68.4 | 9 | 1 | US-08-465-167A-33 |
| 24 | 38 | 66.7 | 591 | 4 | US-09-540-236-3375 |
| 25 | 37 | 64.9 | 9 | 1 | US-08-465-167A-31 |
| 26 | 37 | 64.9 | 10 | 1 | US-08-465-167A-39 |
| 27 | 37 | 64.9 | 10 | 1 | US-08-465-167A-46 |
| | | | | | Sequence 1, Appli |
| | | | | | Sequence 1, Appli |
| | | | | | Sequence 24, Appl |
| | | | | | Sequence 10, Appl |
| | | | | | Sequence 10, Appl |
| | | | | | Sequence 24, Appl |
| | | | | | Sequence 10, Appl |
| | | | | | Sequence 29, Appl |
| | | | | | Sequence 15, Appl |
| | | | | | Sequence 595, App |
| | | | | | Sequence 15, Appl |
| | | | | | Sequence 49, Appl |
| | | | | | Sequence 8, Appli |
| | | | | | Sequence 27, Appl |
| | | | | | Sequence 93, Appl |
| | | | | | Sequence 152, App |
| | | | | | Sequence 8, Appli |
| | | | | | Sequence 44, Appl |
| | | | | | Sequence 93, Appl |
| | | | | | Sequence 152, App |
| | | | | | Sequence 5371, Ap |
| | | | | | Sequence 33, Appl |
| | | | | | Sequence 3375, Ap |
| | | | | | Sequence 31, Appl |
| | | | | | Sequence 39, Appl |
| | | | | | Sequence 46, Appl |

| | | | | | | |
|----|----|------|-----|---|----------------------|-------------------|
| 28 | 35 | 61.4 | 589 | 4 | US-09-107-532A-6086 | Sequence 6086, Ap |
| 29 | 35 | 61.4 | 656 | 4 | US-09-252-991A-29575 | Sequence 29575, A |
| 30 | 34 | 59.6 | 213 | 4 | US-09-328-352-7428 | Sequence 7428, Ap |
| 31 | 34 | 59.6 | 349 | 4 | US-09-252-991A-27115 | Sequence 27115, A |
| 32 | 34 | 59.6 | 378 | 4 | US-09-107-532A-4777 | Sequence 4777, Ap |
| 33 | 32 | 56.1 | 9 | 1 | US-08-465-167A-36 | Sequence 36, Appl |
| 34 | 32 | 56.1 | 102 | 4 | US-09-543-681A-4334 | Sequence 4334, Ap |
| 35 | 32 | 56.1 | 127 | 3 | US-08-906-769-170 | Sequence 170, App |
| 36 | 32 | 56.1 | 127 | 3 | US-08-906-616-170 | Sequence 170, App |
| 37 | 32 | 56.1 | 127 | 3 | US-08-639-075A-170 | Sequence 170, App |
| 38 | 32 | 56.1 | 127 | 3 | US-09-012-431-170 | Sequence 170, App |
| 39 | 32 | 56.1 | 127 | 3 | US-09-012-692-170 | Sequence 170, App |
| 40 | 32 | 56.1 | 127 | 3 | US-08-906-613-170 | Sequence 170, App |
| 41 | 32 | 56.1 | 178 | 3 | US-08-906-769-172 | Sequence 172, App |
| 42 | 32 | 56.1 | 178 | 3 | US-08-906-616-172 | Sequence 172, App |
| 43 | 32 | 56.1 | 178 | 3 | US-08-639-075A-172 | Sequence 172, App |
| 44 | 32 | 56.1 | 178 | 3 | US-09-012-431-172 | Sequence 172, App |
| 45 | 32 | 56.1 | 178 | 3 | US-09-012-692-172 | Sequence 172, App |

ALIGNMENTS

RESULT 1
US-08-465-167A-1
; Sequence 1, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-1

Query Match 100.0%; Score 57; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 EYVIKVSARVRF 12
      |||
Db     30 EYVIKVSARVRF 41

RESULT 2
US-08-627-820-1
; Sequence 1, Application US/08627820
; Patent No. 6464980
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; Livingston, Brian D.
; Sette, Alessandro D.
; Sidney, John C.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
; COMPLETE MAGE 1 GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,820
; FILING DATE: 02-Apr-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,623
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-627-820-1
Query Match      100.0%; Score 57; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      |||
Db     30 EYVIKVSARVRF 41

RESULT 3
US-08-465-167A-24
; Sequence 24, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL

```

```

; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-24
Query Match      100.0%; Score 57; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      |||
Db     281 EYVIKVSARVRF 292

RESULT 4
US-08-993-118-10
; Sequence 10, Application US/08993118
; Patent No. 5997872
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie;
; APPLICANT: DE SMET, Charles;
; APPLICANT: BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
; REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,118
; FILING DATE:

```



```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; US-08-993-118-10

Query Match 100.0%; Score 57; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
Db 281 EYVIKVSARVRF 292

RESULT 5
US-08-845-528C-10
; Sequence 10, Application US/08845528C
; Patent No. 6027924
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie;
; APPLICANT: DE SMET, Charles;
; APPLICANT: BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
; TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,528C
; FILING DATE: April 25, 1997
; CLASSIFICATION: 4335
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; US-08-845-528C-10

Query Match 100.0%; Score 57; DB 3; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0026;

```

```

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
Db 281 EYVIKVSARVRF 292

RESULT 6
US-08-627-820-24
; Sequence 24, Application US/08627820
; Patent No. 6464980
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
; TITLE OF INVENTION: COMPLETE MAGE 1 GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,820
; FILING DATE: 02-Apr-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,623
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-08-627-820-24

Query Match 100.0%; Score 57; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
Db 281 EYVIKVSARVRF 292

RESULT 7
US-09-066-281B-10
; Sequence 10, Application US/09066281B
; Patent No. 6475783
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; DE SMET, Charles; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING
; TITLE OF INVENTION: FOR TUMOR REJECTION ANTIGEN PRECURSOR MAGE-C1 AND MAGE-C2
; TITLE OF INVENTION: AND USES THEREOF

```

```
;
;
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,281B
; FILING DATE: April 24, 1998
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455.2 US - JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
; US-09-066-281B-10
;
; Query Match 100.0%; Score 57; DB 4; Length 309;
; Best Local Similarity 100.0%; Pred. No. 0.0026;
; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 EYVIKVSARVRF 12
; | | | | | | | | | |
; Db 281 EYVIKVSARVRF 292
;
; RESULT 8
; US-09-468-433C-10
; Sequence 10, Application US/09468433C
; Patent No. 6680191
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING FOR
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C AN
; TITLE OF INVENTION: MAGE-B FAMILIES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/468,433C
; FILING DATE: December 17, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/066,281
; FILING DATE: April 24, 1998
```

```
;
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5611 JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 662-0200
; TELEFAX: (202) 662-4643
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
; US-09-468-433C-10
;
; Query Match 100.0%; Score 57; DB 4; Length 309;
; Best Local Similarity 100.0%; Pred. No. 0.0026;
; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 EYVIKVSARVRF 12
; | | | | | | | | | |
; Db 281 EYVIKVSARVRF 292
;
; RESULT 9
; US-09-392-714-29
; Sequence 29, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; US-09-392-714-29
;
; Query Match 100.0%; Score 57; DB 4; Length 309;
; Best Local Similarity 100.0%; Pred. No. 0.0026;
; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 EYVIKVSARVRF 12
; | | | | | | | | | |
; Db 281 EYVIKVSARVRF 292
;
; RESULT 10
; US-08-465-167A-15
; Sequence 15, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
```

```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-167A-15

Query Match      80.7%; Score 46; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 YVIKVSARVR 11
Db      1 YVIKVSARVR 10

RESULT 11
US-08-159-339A-595
; Sequence 595, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

```

;
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-005030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 595:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-159-339A-595

Query Match      80.7%; Score 46; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 YVIKVSARVR 11
Db      1 YVIKVSARVR 10

RESULT 12
US-08-627-820-15
; Sequence 15, Application US/08627820
; Patent No. 6464980
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF THE
; TITLE OF INVENTION: COMPLETE MAGE 1 GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,820
; FILING DATE: 02-Apr-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,623
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids

```



```

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-08-627-820-15

Query Match      80.7%; Score 46; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 YVIKVSARV 11
Db      1 YVIKVSARV 10
      |||||
      |||||

RESULT 13
US-08-465-167A-49
; Sequence 49, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-49

Query Match      73.7%; Score 42; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSAR 9
Db      2 EYVIKVSAR 10
      |||||
      |||||

RESULT 14
US-08-465-167A-49
; Sequence 49, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-49

Query Match      73.7%; Score 42; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSAR 9
Db      2 EYVIKVSAR 10
      |||||
      |||||

RESULT 14
US-08-465-167A-49
; Sequence 49, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-49

Query Match      71.9%; Score 41; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 YVIKVSARV 10
Db      1 YVIKVSARV 9
      |||||
      |||||

RESULT 15
US-08-465-167A-27
; Sequence 27, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-8
```

ZIP: 98111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,167A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/103,623
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-60-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-167A-27

Query Match 71.9%; Score 41; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 YVIKVSARV 10
Db 1 YVIKVSARV 9

Search completed: October 7, 2004, 15:19:26
Job time : 14.0909 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:30:52 ; Search time 44.7273 Seconds
(without alignments)
86.336 Million cell updates/sec

Title: US-09-336-091-7
Perfect score: 57
Sequence: 1 EYVIKVSARVRF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------------------|-------------------|
| 1 | 57 | 100.0 | 12 | 15 US-10-164-121A-33 | Sequence 33, Appl |
| 2 | 57 | 100.0 | 12 | 15 US-10-164-078A-32 | Sequence 32, Appl |
| 3 | 57 | 100.0 | 30 | 12 US-10-296-734-1278 | Sequence 1278, Ap |
| 4 | 57 | 100.0 | 309 | 9 US-09-766-889A-2 | Sequence 2, Appli |
| 5 | 57 | 100.0 | 309 | 12 US-10-296-734-828 | Sequence 828, App |
| 6 | 57 | 100.0 | 309 | 13 US-10-085-108-10 | Sequence 10, Appl |
| 7 | 57 | 100.0 | 309 | 14 US-10-177-390-18 | Sequence 18, Appl |
| 8 | 57 | 100.0 | 309 | 14 US-10-160-237-10 | Sequence 10, Appl |
| 9 | 57 | 100.0 | 309 | 15 US-10-117-937-71 | Sequence 71, Appl |
| 10 | 57 | 100.0 | 309 | 16 US-10-657-022-71 | Sequence 71, Appl |
| 11 | 57 | 100.0 | 309 | 16 US-10-741-466-6 | Sequence 6, Appli |
| 12 | 57 | 100.0 | 311 | 16 US-10-741-466-7 | Sequence 7, Appli |
| 13 | 57 | 100.0 | 311 | 16 US-10-741-466-8 | Sequence 8, Appli |
| 14 | 57 | 100.0 | 3541 | 12 US-10-296-734-1454 | Sequence 1454, Ap |
| 15 | 46 | 80.7 | 10 | 15 US-10-117-937-129 | Sequence 129, App |

| | | | | | | |
|----|----|------|------|----|----------------------|-------------------|
| 16 | 41 | 71.9 | 9 | 12 | US-10-149-135-2182 | Sequence 2182, Ap |
| 17 | 41 | 71.9 | 9 | 12 | US-09-935-476-6 | Sequence 6, Appli |
| 18 | 41 | 71.9 | 9 | 14 | US-10-128-711-93 | Sequence 93, Appl |
| 19 | 41 | 71.9 | 9 | 14 | US-10-128-711-152 | Sequence 152, App |
| 20 | 41 | 71.9 | 9 | 14 | US-10-150-797-13 | Sequence 13, Appl |
| 21 | 41 | 71.9 | 9 | 15 | US-10-149-138-4195 | Sequence 4195, Ap |
| 22 | 41 | 71.9 | 9 | 16 | US-10-149-138-4195 | Sequence 4195, Ap |
| 23 | 41 | 71.9 | 28 | 12 | US-10-296-734-1280 | Sequence 1280, Ap |
| 24 | 41 | 71.9 | 592 | 12 | US-10-282-122A-44745 | Sequence 44745, A |
| 25 | 39 | 68.4 | 9 | 15 | US-10-117-937-128 | Sequence 128, App |
| 26 | 39 | 68.4 | 317 | 12 | US-10-218-095-2 | Sequence 2, Appli |
| 27 | 39 | 68.4 | 317 | 14 | US-10-157-031-52 | Sequence 52, Appl |
| 28 | 38 | 66.7 | 597 | 12 | US-10-282-122A-63003 | Sequence 63003, A |
| 29 | 37 | 64.9 | 226 | 12 | US-10-424-599-274271 | Sequence 274271, |
| 30 | 37 | 64.9 | 543 | 12 | US-10-282-122A-70042 | Sequence 70042, A |
| 31 | 37 | 64.9 | 1887 | 15 | US-10-369-493-22455 | Sequence 22455, A |
| 32 | 37 | 64.9 | 1887 | 16 | US-10-741-191-19 | Sequence 19, Appl |
| 33 | 37 | 64.9 | 1887 | 16 | US-10-742-350-19 | Sequence 19, Appl |
| 34 | 36 | 63.2 | 176 | 14 | US-10-156-761-9282 | Sequence 9282, Ap |
| 35 | 36 | 63.2 | 602 | 12 | US-10-282-122A-66049 | Sequence 66049, A |
| 36 | 36 | 63.2 | 1070 | 7 | US-08-908-453-11 | Sequence 11, Appl |
| 37 | 36 | 63.2 | 1070 | 12 | US-10-170-385-305 | Sequence 305, App |
| 38 | 35 | 61.4 | 47 | 12 | US-10-424-599-276698 | Sequence 276698, |
| 39 | 35 | 61.4 | 269 | 12 | US-10-243-552-530 | Sequence 530, App |
| 40 | 35 | 61.4 | 383 | 12 | US-10-424-599-282528 | Sequence 282528, |
| 41 | 35 | 61.4 | 589 | 12 | US-10-282-122A-58041 | Sequence 58041, A |
| 42 | 35 | 61.4 | 591 | 9 | US-09-815-242-11815 | Sequence 11815, A |
| 43 | 35 | 61.4 | 591 | 12 | US-10-282-122A-66236 | Sequence 66236, A |
| 44 | 35 | 61.4 | 591 | 12 | US-10-282-122A-68348 | Sequence 68348, A |
| 45 | 34 | 59.6 | 45 | 12 | US-10-424-599-168734 | Sequence 168734, |

• ALIGNMENTS

RESULT 1

US-10-164-121A-33
; Sequence 33, Application US/10164121A
; Publication No. US20030228308A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Yi
; APPLICANT: Boon, Thierry
; APPLICANT: Van der Bruggen, Pierre
; APPLICANT: Traversari, Catra
; TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-Cw6 Molecules And Uses There
; FILE REFERENCE: LUD-5771
; CURRENT APPLICATION NUMBER: US/10/164,121A
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 33
; LENGTH: 12
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
US-10-164-121A-33

Query Match 100.0%; Score 57; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EYVIKVSARVRF 12
| | | | | | | | | |
Db 1 EYVIKVSARVRF 12

RESULT 2

US-10-164-078A-32
; Sequence 32, Application US/10164078A
; Publication No. US20030228325A1
; GENERAL INFORMATION:
; APPLICANT: Bilsborough, Janine
; APPLICANT: Schultz, Erwin
; APPLICANT: Panichelli, Christophe

```

; APPLICANT: Van der Bruggen, Pierre
; APPLICANT: Boon, Thierry
; TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-B18 Molecules And Uses Thereof
; FILE REFERENCE: LUD-5756
; CURRENT APPLICATION NUMBER: US/10/164,078A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 32
; LENGTH: 12
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
US-10-164-078A-32

Query Match      100.0%; Score 57; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
Db      1 EYVIKVSARVRF 12

RESULT 3
US-10-296-734-1278
; Sequence 1278, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1278
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MAGE-1 segment 19
US-10-296-734-1278

Query Match      100.0%; Score 57; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00058;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
Db      13 EYVIKVSARVRF 24

RESULT 4
US-09-766-889A-2
; Sequence 2, Application US/09766889A
; Patent No. US20020164654A1
; GENERAL INFORMATION:
; APPLICANT: Luiten, Rosalie
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: van der Bruggen, Pierre
; APPLICANT: Stroobant, Vincent
; APPLICANT: Demotte, Nathalie
; APPLICANT: Schultz, Erwin
; TITLE OF INVENTION: MAGE ANTIGENIC PEPTIDES WHICH BIND HLA-B35 AND HLA-B44
; FILE REFERENCE: L0461/7104
; CURRENT APPLICATION NUMBER: US/09/766,889A
; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/177,242
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/243,212

```

```

; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-766-889A-2

Query Match      100.0%; Score 57; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
Db      281 EYVIKVSARVRF 292

RESULT 5
US-10-296-734-828
; Sequence 828, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 828
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MAGE-1 consensus polypeptide
US-10-296-734-828

Query Match      100.0%; Score 57; DB 12; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
Db      281 EYVIKVSARVRF 292

RESULT 6
US-10-085-108-10
; Sequence 10, Application US/10085108
; Publication No. US20020176865A1
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING FOR
; TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C
; MAGE-B FAMILIES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/085,108
; FILING DATE: 01-Mar-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/501,104
; FILING DATE: 09-Feb-2000
; APPLICATION NUMBER: 09/468,433
; FILING DATE: December 17, 1999
; APPLICATION NUMBER: 09/066,281
; FILING DATE: April 24, 1998
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997.
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5611.1 JEL/MAS
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 318-3400
;
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-085-108-10

Query Match 100.0%; Score 57; DB 13; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
| | | | | | | | | | | |
Db 281 EYVIKVSARVRF 292

RESULT 7
US-10-177-390-18
; Sequence 18, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerold
; APPLICANT: N.V. Antwerps Innovatiecentrum
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with Linear
; TITLE OF INVENTION: Polynucleotides by Electroporation
; FILE REFERENCE: 021505wo/JH/ml
; CURRENT APPLICATION NUMBER: US/10/177,390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-18

Query Match 100.0%; Score 57; DB 14; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
| | | | | | | | | | | |
Db 281 EYVIKVSARVRF 292

RESULT 8
US-10-160-237-10
; Sequence 10, Application US/10160237
; Publication No. US20030170256A1
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; DE SMET, Charles; BOON-FALLEUR, Thierry

```

```

;
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING
; FOR TUMOR REJECTION ANTIGEN PRECURSOR MAGE-C1 AND MAGE-C2
; AND USES THEREOF
;
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/160,237
; FILING DATE: 04-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,281B
; FILING DATE: April 24, 1998
; APPLICATION NUMBER: 08/845,528
; FILING DATE: April 25, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5455.2 US - JEL/MAS
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 752-5958
;
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309
; TYPE: amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-160-237-10

Query Match 100.0%; Score 57; DB 14; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
| | | | | | | | | | | |
Db 281 EYVIKVSARVRF 292

RESULT 9
US-10-117-937-71
; Sequence 71, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: CTLM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71

```



```

;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-117-937-71

Query Match      100.0%; Score 57; DB 15; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      ||| ||| ||| ||| |||
Db      281 EYVIKVSARVRF 292

RESULT 10
US-10-657-022-71
; Sequence 71, Application US/10657022
; Publication No. US20040180354A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Liu, Liping
; APPLICANT: Liu, Zheng
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: MANNK.032A
; CURRENT APPLICATION NUMBER: US/10/657,022
; CURRENT FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: 60/409123
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 610
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-657-022-71

Query Match      100.0%; Score 57; DB 16; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      ||| ||| ||| ||| |||
Db      281 EYVIKVSARVRF 292

RESULT 11
US-10-741-466-6
; Sequence 6, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Influenza A virus
US-10-741-466-6

Query Match      100.0%; Score 57; DB 16; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      ||| ||| ||| ||| |||
Db      281 EYVIKVSARVRF 292

RESULT 12
US-10-741-466-7
; Sequence 7, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Influenza A virus
US-10-741-466-7

Query Match      100.0%; Score 57; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      ||| ||| ||| ||| |||
Db      283 EYVIKVSARVRF 294

RESULT 13
US-10-741-466-8
; Sequence 8, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Influenza A virus
US-10-741-466-8

Query Match      100.0%; Score 57; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      ||| ||| ||| ||| |||
Db      283 EYVIKVSARVRF 294

RESULT 14
US-10-296-734-1454
; Sequence 1454, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
```

```

Db      281 EYVIKVSARVRF 292

RESULT 12
US-10-741-466-7
; Sequence 7, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Influenza A virus
US-10-741-466-7

Query Match      100.0%; Score 57; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      ||| ||| ||| ||| |||
Db      283 EYVIKVSARVRF 294

RESULT 13
US-10-741-466-8
; Sequence 8, Application US/10741466
; Publication No. US20040180058A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, M.
; APPLICANT: Shneider, A.
; TITLE OF INVENTION: Vaccine Compositions and Methods
; FILE REFERENCE: 25955-003
; CURRENT APPLICATION NUMBER: US/10/741,466
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 60/435,500
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Influenza A virus
US-10-741-466-8

Query Match      100.0%; Score 57; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYVIKVSARVRF 12
      ||| ||| ||| ||| |||
Db      283 EYVIKVSARVRF 294

RESULT 14
US-10-296-734-1454
; Sequence 1454, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
```

; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1454
; LENGTH: 3541
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Melanoma cancer specific savine
US-10-296-734-1454

Query Match 100.0%; Score 57; DB 12; Length 3541;
Best Local Similarity 100.0%; Pred. No. 0.095;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYIKVSARVRF 12
| | | | | | | | | |
Db 1704 EYIKVSARVRF 1715

RESULT 15
US-10-117-937-129
; Sequence 129, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: CTLLMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-937-129

Query Match 80.7%; Score 46; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVIVKSARVR 11
| | | | | | | |
Db 1 YVIVKSARVR 10

Search completed: October 7, 2004, 15:33:47
Job time : 44.7273 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:17:11 ; Search time 13.9091 Seconds
(without alignments)
82.989 Million cell updates/sec

Title: US-09-336-091-7
Perfect score: 57
Sequence: 1 EYVIKVSARVRF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 39 | 68.4 | 317 | 2 I38661 | melanoma antigen |
| 2 | 37 | 64.9 | 357 | 2 H71122 | hypothetical prote |
| 3 | 37 | 64.9 | 1887 | 2 S61703 | fatty-acid synthas |
| 4 | 36 | 63.2 | 392 | 2 S72753 | hypothetical prote |
| 5 | 36 | 63.2 | 594 | 2 F72265 | conserved hypothet |
| 6 | 36 | 63.2 | 602 | 2 G81195 | aspartyl-trNA synt |
| 7 | 36 | 63.2 | 602 | 2 F81831 | aspartate-trNA lig |
| 8 | 36 | 63.2 | 1070 | 1 A54600 | 1-phosphatidylinos |
| 9 | 35 | 61.4 | 591 | 2 E83524 | aspartyl-trNA synt |
| 10 | 34 | 59.6 | 255 | 2 D83036 | hypothetical prote |
| 11 | 34 | 59.6 | 288 | 2 AE2832 | conserved hypothet |
| 12 | 34 | 59.6 | 317 | 2 A97610 | hypothetical prote |
| 13 | 34 | 59.6 | 493 | 2 H70479 | leucine aminopepti |
| 14 | 34 | 59.6 | 534 | 2 S57974 | hypothetical prote |
| 15 | 34 | 59.6 | 556 | 2 H84721 | probable preprotei |
| 16 | 34 | 59.6 | 26926 | 1 I38344 | titin, cardiac mus |
| 17 | 33.5 | 58.8 | 497 | 2 S19343 | probable iron-sulf |
| 18 | 33 | 57.9 | 118 | 2 D96788 | protein T4012.19 [|
| 19 | 33 | 57.9 | 323 | 2 S56777 | probable membrane |
| 20 | 33 | 57.9 | 355 | 2 T24822 | hypothetical prote |
| 21 | 33 | 57.9 | 450 | 2 C75316 | conserved hypothet |
| 22 | 33 | 57.9 | 538 | 2 G69317 | conserved hypothet |
| 23 | 33 | 57.9 | 595 | 2 AE2110 | aspartate-trNA lig |
| 24 | 33 | 57.9 | 649 | 1 TVVPMK | large T antigen - |
| 25 | 33 | 57.9 | 846 | 2 H70871 | hypothetical prote |
| 26 | 33 | 57.9 | 869 | 2 A86983 | conserved hypothet |
| 27 | 33 | 57.9 | 869 | 2 S72760 | ppsl protein - Myc |
| 28 | 33 | 57.9 | 946 | 2 T01460 | hypothetical prote |
| 29 | 33 | 57.9 | 1225 | 2 B95229 | DNA-directed RNA p |

| | | | | | |
|----|----|------|------|----------|--------------------|
| 30 | 33 | 57.9 | 1225 | 2 F98093 | DNA-directed RNA p |
| 31 | 33 | 57.9 | 1253 | 2 F86436 | hypothetical prote |
| 32 | 33 | 57.9 | 1357 | 2 T16860 | hypothetical prote |
| 33 | 33 | 57.9 | 1885 | 1 JC4086 | fatty-acid synthas |
| 34 | 33 | 57.9 | 2944 | 2 A54849 | collagen alpha 1(V |
| 35 | 33 | 57.9 | 6805 | 2 S20901 | titin - rabbit (fr |
| 36 | 32 | 56.1 | 33 | 2 E72309 | hypothetical prote |
| 37 | 32 | 56.1 | 126 | 2 F82611 | transcription regu |
| 38 | 32 | 56.1 | 204 | 2 S12121 | rho protein GDP-di |
| 39 | 32 | 56.1 | 204 | 2 I38156 | rho protein GDP-di |
| 40 | 32 | 56.1 | 208 | 2 T16138 | histone H1.3 - Cae |
| 41 | 32 | 56.1 | 230 | 1 JN0060 | hypothetical 24.5K |
| 42 | 32 | 56.1 | 263 | 2 T01614 | hypothetical prote |
| 43 | 32 | 56.1 | 270 | 2 E69303 | methyltransferase |
| 44 | 32 | 56.1 | 278 | 2 F71036 | probable dipeptide |
| 45 | 32 | 56.1 | 318 | 2 E84360 | succinate dehydrog |

ALIGNMENTS

RESULT 1

I38661
melanoma antigen MAGE-4 - human
N;Alternate names: MAGE 41 protein; melanoma antigen MAGE-X2
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1996 #sequence revision 07-Jun-1996 #text change 18-Feb-2000
C;Accession: I38661; I38662; PH1297; PH1298; JC2359; G01446
R;De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; B
oon, T.
Immunogenetics 40, 360-369, 1994
A;Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE fa
A;Reference number: I38659; MUID:95012457; PMID:7927540
A;Accession: I38661
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-317 <DEP1>
A;Cross-references: EMBL:U10687; NID:G533514; PIDN:AAA68871.1; PID:G533515
A;Experimental source: antigen MAGE-4a
A;Accession: I38662
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-172,'T',174-317 <DEP2>
A;Cross-references: EMBL:U10688; NID:G533516; PIDN:AAA68872.1; PID:G533517
A;Experimental source: antigen MAGE-4b
R;Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van Pel
J. Exp. Med. 176, 1453-1457, 1992
A;Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolyti
A;Reference number: PH1294; MUID:93018875; PMID:1402688
A;Accession: PH1297
A;Molecule type: DNA
A;Residues: 169-177 <TRAI>
A;Experimental source: antigen MAGE-4
A;Accession: PH1298
A;Molecule type: DNA
A;Residues: 169-172,'T',174-177 <TRA2>
A;Experimental source: antigen MAGE-41
R;Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
Biochem. Biophys. Res. Commun. 202, 549-555, 1994
A;Title: Cloning and analysis of MAGE-1-related genes.
A;Reference number: JC2358; MUID:94311935; PMID:8037761
A;Accession: JC2359
A;Molecule type: mRNA
A;Residues: 1-172,'T',174-306,'Q',308-317 <DIN>
A;Cross-references: EMBL:U10340; NID:G499123; PIDN:AAA19007.1; PID:G499124
A;Experimental source: melanoma cell line DM150
C;Genetics:
A;Gene: GDB:MAGEA4; MAGE4; MAGE-X2
A;Cross-references: GDB:331119
A;Map position: Xq28-Xq28
A;Introns: #status absent
C;Superfamily: tumor associated protein MAGE
F;169-177/Region: HLA-A1 binding #status predicted

Query Match 68.4%; Score 39; DB 2; Length 317;
Best Local Similarity 63.6%; Pred. No. 3.3;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|:|:|:|

Db 289 EHVVRVNARVR 299

RESULT 2
H71122
hypothetical protein PH0751 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 02-Feb-2001
C;Accession: H71122
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: H71122
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <KAW>
A;Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29842.1; PID:g3257159
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0751
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 64.9%; Score 37; DB 2; Length 357;
Best Local Similarity 72.7%; Pred. No. 9.9;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|:|:|:|

Db 297 EYPIEVSGRVR 307

RESULT 3
S61703
fatty-acid synthase (EC 2.3.1.85) alpha chain - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein P1409; protein YPL231w
C;Species: Saccharomyces cerevisiae
C;Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 20-Jun-2000
C;Accession: S61703; S65256; S65250; A31107
R;Urrestarazu, L.A.
submitted to the EMBL Data Library, December 1995
A;Reference number: S61699
A;Accession: S61703
A;Molecule type: DNA
A;Residues: 1-1887 <URR>
A;Cross-references: EMBL:X94561; NID:g1181252; PIDN:CAA64256.1; PID:g1181257
R;Urrestarazu, L.A.; Visser, S.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S65251
A;Accession: S65256
A;Molecule type: DNA
A;Residues: 1-1887 <URW>
A;Cross-references: EMBL:Z73587; NID:g1370477; PIDN:CAA97948.1; PID:g1370478; MIPS:YPL23
A;Experimental source: strain S288C (AB972)
R;Rieger, M.; Mueller-Auer, S.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S65202
A;Accession: S65250
A;Molecule type: DNA
A;Residues: 1567-1887 <RIE>
A;Cross-references: EMBL:Z73587; MIPS:YPL231w
A;Experimental source: strain S288C (AB972)
R;Mohamed, A.H.; Chirala, S.S.; Mody, N.H.; Huang, W.Y.; Wakil, S.J.
J. Biol. Chem. 263, 12315-12325, 1988

A;Title: Primary structure of the multifunctional alpha subunit protein of yeast fatty
A;Reference number: A31107; MUID:88315020; PMID:2900835
A;Accession: A31107
A;Molecule type: DNA
A;Residues: 1-310, 'TTGTGG', 311-593, 'I', 595-940, 'CLNCVKSWLKLLKLERQFPSKLLW', 'SIRLSMAIALMLI
A;Cross-references: EMBL:J03936; NID:g171501; PIDN:AAA34601.1; PID:g171502
C;Genetics:
A;Gene: SGD:FAS2
A;Cross-references: SGD:S0006152; MIPS:YPL231w
A;Map position: 16L
C;Superfamily: yeast fatty-acid synthase
C;Keywords: acyltransferase; coenzyme A

Query Match 64.9%; Score 37; DB 2; Length 1887;
Best Local Similarity 88.9%; Pred. No. 55;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYVIKVSAR 9
|:|:|:|

Db 1672 EYVAKVSAR 1680

RESULT 4
S72753
hypothetical protein B1496_C1_154 - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001
C;Accession: S72753; T11012
R;Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A;Description: Mycobacterium leprae cosmid B1496.
A;Reference number: S72695
A;Accession: S72753
A;Molecule type: DNA
A;Residues: 1-392 <SMI>
A;Cross-references: EMBL:U00013; NID:g466868; PIDN:AAA17119.1; PID:g466874
R;Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z16918
A;Accession: T11012
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-392 <PAR>
A;Cross-references: EMBL:Z99125; NID:g23398683; PIDN:CAB16170.1; PID:e343547; PID:g233987
C;Genetics:
A;Gene: MLC1536.27C

Query Match 63.2%; Score 36; DB 2; Length 392;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VIKVSARVRF 12
|:|:|:|

Db 208 VVRVSANVRF 217

RESULT 5
F72265
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72265
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: F72265
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-594 <ARN>
A;Cross-references: GB:AE001788; GB:AE000512; NID:g4981893; PIDN:AAD36406.1; PID:g49818

A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1334

Query Match 63.2%; Score 36; DB 2; Length 594;
Best Local Similarity 63.6%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||:|:|
Db 289 EYIIKGGRRV 299

RESULT 6
G81195
aspartyl-tRNA synthetase NMB0466 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81195
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, A.; Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81195
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-602 <TET>
A;Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40903.1; PID:g722569

A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0466
C;Superfamily: lysine-tRNA ligase

Query Match 63.2%; Score 36; DB 2; Length 602;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||:|:|
Db 66 EYVLSITGRV 76

RESULT 7
F81831
aspartate-tRNA ligase (EC 6.1.1.12) NMA2019 [imported] - Neisseria meningitidis (strain F81831)
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C;Accession: F81831
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, A.; Stover, C.K.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, A.; Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: F81831
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-602 <PAR>
A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85238.1; PID:g738064

A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: asps; NMA2019
C;Superfamily: lysine-tRNA ligase
C;Keywords: ligase

Query Match 63.2%; Score 36; DB 2; Length 602;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||:|:|

Db 66 EYVLSITGRV 76

RESULT 8

A54600
1-phosphatidylinositol 3-kinase (EC 2.7.1.137) 110K chain beta isoform - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A54600
R;Hu, P.; Mondino, A.; Skolnik, E.Y.; Schlessinger, J. Mol. Cell. Biol. 13, 7677-7688, 1993
A;Title: Cloning of a novel, ubiquitously expressed human phosphatidylinositol 3-kinase
A;Reference number: A54600; MUID:94067128; PMID:8246984
A;Accession: A54600
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1070 <HU1>
A;Cross-references: GB:S67334; NID:g455759; PIDN:AAB29081.1; PID:g455760
A;Note: sequence extracted from NCBI backbone (NCBIN:140879, NCBIPI:140880)
C;Genetics:
A;Gene: GDB:PIK3CB; PIK3C1
A;Cross-references: GDB:I36233
C;Superfamily: phosphatidylinositol 3-kinase
C;Keywords: phosphotransferase

Query Match 63.2%; Score 36; DB 1; Length 1070;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 12
|||:|:|
Db 245 DYVLQVSGRVEY 256

RESULT 9

E83524
aspartyl-tRNA synthetase PA0963 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83524
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, A.; Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83524
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <STO>
A;Cross-references: GB:AE004530; GB:AE004091; NID:g9946865; PIDN:AAG04352.1; GSPDB:GN00

A;Experimental source: strain PA01
C;Genetics:
A;Gene: asps; PA0963
C;Superfamily: lysine-tRNA ligase

Query Match 61.4%; Score 35; DB 2; Length 591;
Best Local Similarity 45.5%; Pred. No. 43;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||:|:|
Db 67 EFVVKITGKVR 77

RESULT 10

D83036
hypothetical protein PA4882 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: D83036
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim


```
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83036
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-255 <STO>
A;Cross-references: GB:AE004901; GB:AE004091; NID:g9951147; PIDN:AAG08267.1; GSPDB:GN001
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4882

    Query Match      59.6%; Score 34; DB 2; Length 255;
    Best Local Similarity 63.6%; Pred. No. 30;
    Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVR 11
    ||:|||||
Db 108 EYRLQVSGRVR 118

RESULT 11
AE2832
conserved hypothetical protein Atu2084 [imported] - Agrobacterium tumefaciens (strain C5
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AE2832
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthrner, D.; Kutyaivin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE2832
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-288 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL43075.1; PID:g17740544; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2084
A;Map position: circular chromosome

    Query Match      59.6%; Score 34; DB 2; Length 288;
    Best Local Similarity 50.0%; Pred. No. 33;
    Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVR 12
    ||| ||:|
Db 166 EYVADAQAQIR 177

RESULT 12
A97610
hypothetical protein AGR_C_3779 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: A97610
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A97610
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-317 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK87834.1; PID:g15157214; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_3779
```

```
A;Map position: circular chromosome

    Query Match      59.6%; Score 34; DB 2; Length 317;
    Best Local Similarity 50.0%; Pred. No. 37;
    Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVR 12
    ||| ||:|
Db 195 EYVADAQAQIR 206

RESULT 13
H70479
leucine aminopeptidase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 03-Mar-2003
C;Accession: H70479
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: H70479
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-493 <AQF>
A;Cross-references: GB:AE000772; NID:g2984299; PIDN:AAC07829.1; PID:g2984306; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: pepA
C;Superfamily: Cytosol aminopeptidase

    Query Match      59.6%; Score 34; DB 2; Length 493;
    Best Local Similarity 60.0%; Pred. No. 58;
    Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EYVIKVSARV 10
    ||:||||:|
Db 484 EYIMKVSSNV 493

RESULT 14
S57974
hypothetical protein YDR147w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YD8358.04
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
C;Accession: S57974
R;Murphy, L.; Richards, C.; Harris, D.
submitted to the EMBL Data Library, July 1995
A;Reference number: S57971
A;Accession: S57974
A;Molecule type: DNA
A;Residues: 1-534 <MUR>
A;Cross-references: EMBL:Z50046; NID:g899393; PID:g899397; GSPDB:GN00004; MIPS:YDR147w
A;Experimental source: strain AB972
C;Genetics:
A;Gene: SGD:EKI1; MIPS:YDR147w
A;Cross-references: SGD:S0002554
A;Map position: 4R

    Query Match      59.6%; Score 34; DB 2; Length 534;
    Best Local Similarity 58.3%; Pred. No. 63;
    Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVR 12
    ||:||||:|
Db 175 EYELKVIARLSF 186

RESULT 15
H84721
probable preprotein translocase SECY protein [imported] - Arabidopsis thaliana
```

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84721
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84721
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-556 <STO>
A;Cross-references: GB:AE002093; NID:g4582448; PIDN:AAD24832.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g31530
A;Map position: 2

Query Match 59.6%; Score 34; DB 2; Length 556;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EYVTKVSARVRF 12
||:|:|
Db 479 EYLTQIASTRF 490

Search completed: October 7, 2004, 15:18:18
Job time : 14.9091 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:13:16 ; Search time 8.18182 Seconds
(without alignments)
76.370 Million cell updates/sec

Title: US-09-336-091-7
Perfect score: 57
Sequence: 1 EYVIKVSARVF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 57 | 100.0 | 309 | 1 MAG1_HUMAN | P43355 homo sapien |
| 2 | 39 | 68.4 | 317 | 1 MAG4_HUMAN | P43358 homo sapien |
| 3 | 39 | 68.4 | 1004 | 1 RPOC_OENOE | P95405 oenococcus |
| 4 | 37 | 64.9 | 591 | 1 SYD_PSESM | Q87y31 pseudomonas |
| 5 | 37 | 64.9 | 1887 | 1 FAS2_YEAST | P19097 s fatty aci |
| 6 | 36 | 63.2 | 392 | 1 YE62_MYCLE | Q49682 mycobacteri |
| 7 | 36 | 63.2 | 602 | 1 SYD_NEIMA | Q9jt23 neisseria m |
| 8 | 36 | 63.2 | 602 | 1 SYD_NEIMB | Q9k0u5 neisseria m |
| 9 | 36 | 63.2 | 1070 | 1 P11B_HUMAN | P42338 homo sapien |
| 10 | 36 | 63.2 | 1070 | 1 P11B_RAT | Q9z110 rattus norv |
| 11 | 35 | 61.4 | 591 | 1 SYD_PSEAE | Q51422 pseudomonas |
| 12 | 35 | 61.4 | 591 | 1 SYD_PSEPK | Q88nj4 pseudomonas |
| 13 | 34 | 59.6 | 122 | 1 RL31_CABEL | Q9u332 caenorhabdi |
| 14 | 34 | 59.6 | 269 | 1 EFTS_CANBF | Q7vres candidatus |
| 15 | 34 | 59.6 | 493 | 1 AMPA_AQUAE | O67868 aquifex aeo |
| 16 | 34 | 59.6 | 534 | 1 EK11_YEAST | Q03764 saccharomyc |
| 17 | 34 | 59.6 | 4705 | 1 FAT2_DROME | Q9vw71 drosophila |
| 18 | 33.5 | 58.8 | 497 | 1 NFS1_YEAST | P25374 saccharomyc |
| 19 | 33 | 57.9 | 228 | 1 YNEB_METEX | Q49116 methylobact |
| 20 | 33 | 57.9 | 323 | 1 CTK2_YEAST | P46962 saccharomyc |
| 21 | 33 | 57.9 | 595 | 1 SYD_ANASP | Q8yub6 anabaena sp |
| 22 | 33 | 57.9 | 648 | 1 TALA_POVMK | P24597 mouse polyo |
| 23 | 33 | 57.9 | 846 | 1 YE61_MYCTU | O53152 mycobacteri |
| 24 | 33 | 57.9 | 869 | 1 YE61_MYCLE | Q49689 mycobacteri |
| 25 | 33 | 57.9 | 1885 | 1 FAS2_CANAL | P43098 c fatty aci |
| 26 | 33 | 57.9 | 2944 | 1 CA17_HUMAN | Q02388 homo sapien |
| 27 | 32 | 56.1 | 51 | 1 RL39_AERPE | P59472 aeropyrum p |
| 28 | 32 | 56.1 | 204 | 1 GDIR_BOVIN | P19803 bos taurus |
| 29 | 32 | 56.1 | 204 | 1 GDIR_HUMAN | P52565 homo sapien |
| 30 | 32 | 56.1 | 204 | 1 GDIR_MOUSE | Q99pt1 mus musculu |
| 31 | 32 | 56.1 | 230 | 1 Y394_PSEAE | P24562 pseudomonas |
| 32 | 32 | 56.1 | 326 | 1 MENC_MYCTU | O06419 mycobacteri |
| 33 | 32 | 56.1 | 364 | 1 TRU1_THETN | Q8rah7 thermoanaer |

RESULT 1
MAG1_HUMAN
ID MAG1_HUMAN STANDARD; PRT; 309 AA.
AC P43355; O00346;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Melanoma-associated antigen 1 (MAGE-1 antigen) (Antigen MZ2-E).
GN MAGEA1 OR MAGE1 OR MAGE1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92086861; PubMed=1840703;
RA van der Bruggen P., Traversari C., Chomez P., Lurquin C., de Plaen E.,
van den Eynde B., Knuth A., Boon T.;
RT "A gene encoding an antigen recognized by cytolytic T lymphocytes on
a human melanoma.";
RL Science 254:1643-1647(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=94311935; PubMed=8037761;
RA Ding M., Beck R.J., Keller C.J., Fenton R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL Biochem. Biophys. Res. Commun. 202:549-555(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20314869; PubMed=10854409;
RA Mallon A.M., Platzer M., Bates R., Gloeckner G., Botcherby M.,
Nordsiek G., Strivens M.A., Kioschis P., Dangel A., Cunningham D.,
Straw R., Weston P., Hunter C., Gilbert M., Fernando S., Goodall K.,
Kerry G., Greystrom J.S., Clark D., Goerdes M., Blechschmidt K.,
Rump A., Hinzmänn B., Mundy C.R., Miller W., Poustka A., Herman G.E.,
Rhodes M., Denny P., Rosenthal A., Brown S.D.M.;
RT "Comparative genome sequence analysis of the Bpa/Str region in mouse
and man.";
RL Genome Res. 10:758-775(2000).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT ALA-32.
RA Chen H., Wang L., Mei M., Qin L., Cong X., Xu J., Wei L., Wang Y.,
Chen W.;
RT "The polymorphism of MAGE-1 gene in Chinese people.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP MUTAGENESIS.
RC TISSUE=Blood;
RX MEDLINE=94157413; PubMed=8113684;
RA Gaugler B., van den Eynde B., van der Bruggen P., Romero P.,
Gaforio J.J., de Plaen E., Lethe B., Brasseur F., Boon T.;
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
autologous cytolytic T lymphocytes.";
RL J. Exp. Med. 179:921-930(1994).
RN [6]

ALIGNMENTS

| | | | | | | |
|----|----|------|-----|---|------------|--------------------|
| 34 | 32 | 56.1 | 419 | 1 | SYH PYRAE | Q8zwz1 pyrobaculum |
| 35 | 32 | 56.1 | 482 | 1 | GLGA_CLOPE | Q8xpal clostridium |
| 36 | 32 | 56.1 | 605 | 1 | APM2_YEAST | P38700 saccharomyc |
| 37 | 32 | 56.1 | 700 | 1 | EFG_FASMU | P57938 pasteurella |
| 38 | 32 | 56.1 | 807 | 1 | SYL_MYCPU | Q98rb6 mycoplasma |
| 39 | 32 | 56.1 | 966 | 1 | AMPN_HUMAN | P15144 homo sapien |
| 40 | 32 | 56.1 | 970 | 1 | TRAI_COMTE | Q04222 comamonas t |
| 41 | 31 | 54.4 | 139 | 1 | RS6_BORBU | O51142 borrelia bu |
| 42 | 31 | 54.4 | 170 | 1 | YF34_AQUAE | O67492 aquifex aeo |
| 43 | 31 | 54.4 | 253 | 1 | KANU_BACSP | P05058 bacillus sp |
| 44 | 31 | 54.4 | 253 | 1 | KANU_STAAU | P05057 staphylococ |
| 45 | 31 | 54.4 | 316 | 1 | E13E_HORVU | Q02438 hordeum vul |

RP SUBCELLULAR LOCATION.
RX MEDLINE=95012905; PubMed=7927954;
RA Schultz-Thater E., Juretic A., Dellabona P., Luscher U., Siegrist W.,
RA Harder F., Heberer M., Zuber M., Spagnoli G.C.;
RT "MAGE-1 gene product is a cytoplasmic protein.";
RL Int. J. Cancer 59:435-439(1994).
CC -!- FUNCTION: Not known, though may play a role in embryonal
CC development and tumor transformation or aspects of tumor
CC progression. Antigen recognized on a melanoma by autologous
CC cytolytic T lymphocytes.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Expressed in many tumors of several types,
CC such as melanoma, head and neck squamous cell carcinoma, lung
CC carcinoma and breast carcinoma, but not in normal tissues except
CC for testes. Never expressed in kidney tumors, leukemias and
CC lymphomas.
CC -!- SIMILARITY: Contains 1 MAGE domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M77481; AAA03229.1; -.
DR EMBL; U82670; -; NOT ANNOTATED_CDS.
DR EMBL; AY148486; AAN62752.1; -.
DR Genew; HGNC:6796; MAGEA1.
DR MIM; 300016; -.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Polymorphism; Tumor antigen.
FT DOMAIN 102 301
FT DOMAIN 33 36
FT VARIANT 32 32
FT VARIANT 72 72
FT MUTAGEN 163 163 T -> A (probable polymorphism).
FT MUTAGEN 169 169 /FTId=VAR_004283.
FT SEQUENCE 309 AA; 34342 MW; 544EEB1F9F4E9D33 CRC64;
Query Match 100.0%; Score 57; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EYVIKVSARVRF 12
Db 281 EYVIKVSARVRF 292
RESULT 2
MAG4_HUMAN
ID MAG4_HUMAN STANDARD; PRT; 317 AA.
AC P43358;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Melanoma-associated antigen 4 (MAGE-4 antigen) (MAGE-X2) (MAGE-41).
GN MAGEA4 OR MAGE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95012457; PubMed=7927540;
RA de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,

RA de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL Immunogenetics 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=94311935; PubMed=8037761;
RA Ding M., Beck R.J., Keller C.J., Penton R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL Biochem. Biophys. Res. Commun. 202:549-555(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95369706; PubMed=7642112;
RA Imai Y., Shichijo S., Yamada A., Katayama T., Yano H., Itoh K.;
RT "Sequence analysis of the MAGE gene family encoding human tumor-
RT rejection antigens.";
RL Gene 160:287-290(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Duodenum;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: Contains 1 MAGE domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10687; AAA68871.1; -.
DR EMBL; U10688; AAA68872.1; -.
DR EMBL; U10340; AAA19007.1; -.
DR EMBL; D32077; BAA06843.1; -.
DR EMBL; BC017723; AAH17723.1; -.
DR PIR; I38661; I38661.
DR PDB; 1I4F; 25-JUL-01.
DR Genew; HGNC:6802; MAGEA4.
DR MIM; 300175; -.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Polymorphism; Tumor antigen; 3D-structure.


```
FT DOMAIN 110 309 MAGE.
FT DOMAIN 41 44 POLY-SER.
FT VARIANT 173 173 T -> A.
FT CONFLICT 307 307 /FTId=VAR 004284.
SQ SEQUENCE 317 AA; 34929 MW; 3CFAC0E2B696257C CRC64;

Query Match 68.4%; Score 39; DB 1; Length 317;
Best Local Similarity 63.6%; Pred. No. 2;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
Db 289 EHVVRNARVR 299

RESULT 3
RPOC_OENOE STANDARD; PRT; 1004 AA.
AC P95405;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
DE beta' chain) (RNA polymerase beta' subunit) (Fragment).
GN RPOC.
OS Oenococcus oeni (Leuconostoc oenos).
OC Bacteria; Firmicutes; Lactobacillales; Oenococcus.
OX NCBI_TaxID=1247;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCDO 1674;
RX MEDLINE=97016803; PubMed=8863429;
RA Morse R., Collins M.D., O'Hanlon K., Wallbanks S., Richardson P.T.;
RT "Analysis of the beta' subunit of DNA-dependent RNA polymerase does
RT not support the hypothesis inferred from 16S rRNA analysis that
RT Oenococcus oeni (formerly Leuconostoc oenos) is a tachytelic
RT (fast-evolving) bacterium."
RL Int. J. Syst. Bacteriol. 46:1004-1009(1996).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: The enzyme consists of the sigma chain and the core
CC enzyme which is composed of 2 alpha chains, 1 beta chain, and 1
CC beta' chain.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X96384; CAA65248.1; -.
CC HSSP; Q9KWU6; 1HQM.
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_pol_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR SMART; SM00663; RPOLA N; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription.
FT NON_TER 1 1
```

```
FT NON_TER 1004 1004
SQ SEQUENCE 1004 AA; 111965 MW; 73750DF47F3A2C36 CRC64;

Query Match 68.4%; Score 39; DB 1; Length 1004;
Best Local Similarity 63.6%; Pred. No. 6.5;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
Db 942 EYVLPISARLR 952

RESULT 4
SYD_PSESM STANDARD; PRT; 591 AA.
AC Q87Y31;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AsPRS).
GN ASPS OR PSPTO3981.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Winn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidse T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,
RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE016870; AAO57440.1; -.
CC TIGR; PSPTO3981; -.
DR HAMAP; MF_00044; -.
DR InterPro; IPR004115; GAD dom.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_asp.
DR InterPro; IPR004365; tRNA_anti.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 1.
DR Pfam; PF01336; tRNA_anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR PROSITE; PS50862; AA TRNA LIGASE II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 591 AA; 66351 MW; 4A841555305F8F36 CRC64;

Query Match 64.9%; Score 37; DB 1; Length 591;
```


Best Local Similarity 63.6%; Pred. No. 9.7;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||:|:| :|||
Db 67 EYVKVVGKVR 77

RESULT 5
FAS2 YEAST STANDARD; PRT; 1887 AA.
AC P19097; Q12533;
DT 01-NOV-1990 (Rel. 16, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fatty acid synthase subunit alpha (EC 2.3.1.86) [Includes: Acyl
carrier; 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
DE (Beta-ketoacyl reductase); 3-oxoacyl-[acyl-carrier protein] synthase
DE (EC 2.3.1.41) (Beta-ketoacyl synthase)].
GN FAS2 OR YPL231W OR P1409.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88315020; PubMed=2900835;
RA Mohamed A.H., Chirala S.S., Mody N.H., Huang W.Y., Wakil S.J.;
RT "Primary structure of the multifunctional alpha subunit protein of
RT yeast fatty acid synthase derived from FAS2 gene sequence.";
RL J. Biol. Chem. 263:12315-12325(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=X2180-1A;
RA Schueller H.J.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albertmann K., Allen E., Ansoerge W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Visser S., Voss H.,
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL Nature 387:103-105(1997).
RN [4]
RP MUTAGENESIS OF GLY-1250.
RX STRAIN=S288C;
RX MEDLINE=94316198; PubMed=8041367;
RA Inokoshi J., Tomoda H., Hashimoto H., Watanabe A., Takeshima H.,
RA Omura S.;
RT "Cerulenin-resistant mutants of Saccharomyces cerevisiae with an
RT altered fatty acid synthase gene.";
RL Mol. Gen. Genet. 244:90-96(1994).
CC -1- FUNCTION: Fatty acid synthetase catalyzes the formation of
CC long-chain fatty acids from acetyl-CoA, malonyl-CoA and NADPH.
CC The alpha subunit contains domains for: acyl carrier protein,
CC 3-oxoacyl-[acyl-carrier protein] reductase, and 3-oxoacyl-[acyl-
CC carrier-protein] synthase. This subunit coordinates the binding
CC of the six beta subunits to the enzyme complex.

CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
CC long-chain acyl-CoA + N CoA + N CO(2) + 2N NADP(+).
CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
CC [acyl-carrier protein].
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1- SUBUNIT: [Alpha(6)beta(6)] hexamers of two multifunctional
CC subunits (alpha and beta).
CC -1- SIMILARITY: TO THE FATTY ACID SYNTHETASE, SUBUNIT ALPHA FROM
CC OTHER FUNGI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC DRL EMBL; J03936; AAA34601.1; -;
CC DRL EMBL; X76890; CAA54218.1; -;
CC DRL EMBL; X94561; CAA64256.1; -;
CC DRL EMBL; Z73586; CAA97947.1; -;
CC DRL EMBL; Z73587; CAA97948.1; -;
CC DRL PIR; S61703; S61703.
CC DRL GermOnline; 144213; -;
CC DRL SGD; S0006152; FAS2.
CC DRL InterPro; IPR008278; 4-PPT_transf.
CC DRL InterPro; IPR000794; Ketoacyl_synth.
CC DRL InterPro; IPR004568; Pantethn_trn.
CC DRL InterPro; IPR006162; Ppantne_S.
CC DRL Pfam; PF01648; ACPS; 1.
CC DRL Pfam; PF00109; ketoacyl-synt; 1.
CC DRL Pfam; PF02801; ketoacyl-synt C; 1.
CC DRL TIGRFAMS; TIGR00556; pantethn_trn; 1.
CC DRL PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
CC DRL PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
CC DRL Fatty acid biosynthesis; Multifunctional enzyme; Oxidoreductase;
CC DRL Transferase; NADP; Phosphopantetheine.
CC DRL DOMAIN 1 ? ACYL CARRIER (ACP).
CC FT DOMAIN 675 874 BETA-KETOACYL REDUCTASE.
CC FT DOMAIN 1149 1363 BETA-KETOACYL SYNTHASE.
CC FT BINDING 180 180 PHOSPHOPANTETHEINE (BY SIMILARITY).
CC FT ACT_SITE 1305 1305 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
CC FT MUTAGEN 1250 1250 G->S: CERULENIN-RESISTANCE.
CC FT CONFLICT 310 310 G -> GTTGTTG (IN REF. 1).
CC FT CONFLICT 594 594 T -> I (IN REF. 1).
CC FT CONFLICT 941 1019 AKRLKELVETSEVRKAVSIETALEHKVVNGNSADAAYAOVE
CC FT IQPRANIQLDFPELPKPKYQVKQIAPAELEGLLDLRVI ->
CC FT CLNCVKSWLKLLKLERQFPKSLMSIRLSMAIALMLHMLKS
CC FT KFNQELTFNWTSONRHNTRLNKLPLSLRVCWIWKELF
CC FT (IN REF. 1).
CC FT CONFLICT 1036 1041 RWMEA -> KMGNGS (IN REF. 1).
CC FT CONFLICT 1408 1408 A -> S (IN REF. 1).
CC FT CONFLICT 1671 1671 N -> T (IN REF. 1).
CC SQ SEQUENCE 1887 AA; 206946 MW; 08B872734CF3AEEA CRC64;

Query Match 64.9%; Score 37; DB 1; Length 1887;
Best Local Similarity 88.9%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYVIKVSAR 9
||| |||||
Db 1672 EYVAKVSAR 1680

RESULT 6
YE62 MYCLE
ID YE62_MYCLE STANDARD; PRT; 392 AA.
AC Q49682;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)

```
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ML0594.
GN ML0594 OR MLC1536.27C OR U1496A OR B1496_C1_154.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -!- SIMILARITY: Belongs to the UPF0051 (ycf24) family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U00013; AAA17119.1; -.
DR EMBL; Z99125; CAB16170.1; -.
DR EMBL; AL583919; CAC30102.1; -.
DR PIR; S72753; S72753.
DR Leproma; ML0594; -.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF01458; UPF0051; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 392 AA; 42202 MW; 891162F7CA494C6A CRC64;

Query Match 63.2%; Score 36; DB 1; Length 392;
Best Local Similarity 70.0%; Pred. No. 10;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VIKVSARVRF 12
Db 208 VVRVSANVRF 217
|:|:|:|:|:|

RESULT 7
SYD_NEIMA
ID SYD_NEIMA STANDARD; PRT; 602 AA.
AC Q9JT23;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (ASPRS).
GN ASPs OR NMA2019.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;

RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AL162757; CAB85238.1; -.
DR PIR; F81831; F81831.
DR HSP; P21889; IEQR.
DR HAMAP; MF_00044; -.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_2.
DR InterPro; IPR004365; tRNA_anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA_anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMs; TIGR00459; asps_bact; 1.
DR PROSITE; PS50862; AA TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 602 AA; 67976 MW; EF41A56255A79491 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 602;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
Db 66 EYVLSITGRVR 76
|:|:|:|:|:|

RESULT 8
SYD_NEIMB
ID SYD_NEIMB STANDARD; PRT; 602 AA.
AC Q9KQ05;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (ASPRS).
GN ASPs OR NMB0466.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
```


RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Massignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
CC -|- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002403; AAF40903.1; -.
CC PIR; G81195; G81195.
CC HSSP; P21889; 1EQR.
CC TIGR; NMB0466; -.
CC HAMAP; MF 00044; -. 1.
CC InterPro; IPR004524; Asps_bact.
CC InterPro; IPR004115; GAD_dom.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC InterPro; IPR004364; tRNA-synt_2.
CC InterPro; IPR002312; tRNA-synt_asp.
CC InterPro; IPR004365; tRNA_anti.
CC InterPro; IPR006195; tRNA_ligase_II.
CC Pfam; PF02938; GAD; 1.
CC Pfam; PF00152; tRNA-synt_2; 2.
CC Pfam; PF01336; tRNA_anti; 1.
CC PRINTS; PR01042; TRNASYNTHASP.
CC TIGRFAMs; TIGR00459; asps_bact; 1.
CC PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 602 AA; 68124 MW; 5DC8A016B0C13E3C CRC64;

Query Match 63.2%; Score 36; DB 1; Length 602;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||: : : |||
Db 66 EYVLSITGRVR 76

RESULT 9
P11B HUMAN STANDARD; PRT; 1070 AA.
AC P42338;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit, beta
DE isoform (EC 2.7.1.153) (PI3-kinase p10 subunit beta) (PtdIns-3-kinase
DE p110) (PI3K) (PI3Kbeta).
GN PIK3CB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94067128; PubMed=8246984;
RA Hu P., Mondino A., Skolnik E.Y., Schlessinger J.;

RT "Cloning of a novel, ubiquitously expressed human
RT phosphatidylinositol 3-kinase and identification of its binding site
RT on p85.";
RL Mol. Cell. Biol. 13:7677-7688(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Kossila M., Sinkovic M., Karkkainen P., Laukkanen M.O., Miettinen R.,
RA Rissanen J., Kekalainen P., Kuusisto J., Yla-Herttuala S., Laakso M.;
RT "Gene encoding the catalytic subunit p110beta of human
RT phosphatidylinositol 3-kinase: cloning, genomic structure and
RT screening for variants in patients with type 2 diabetes.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: PHOSPHORYLATES PTDINS, PTDINS4P AND PTDINS(4,5)P2 WITH A
CC PREFERENCE FOR PTDINS(4,5)P2.
CC -|- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4,5-
CC bisphosphate = ADP + 1-phosphatidyl-1D-myo-inositol 3,4,5-
CC trisphosphate.
CC -|- PATHWAY: Signaling pathways regulating cell growth.
CC -|- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNIT.
CC -|- TISSUE SPECIFICITY: Expressed ubiquitously.
CC -|- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S67334; AAB29081.1; -.
CC EMBL; AJ297549; CAC21449.1; -.
CC EMBL; AJ297550; CAC21449.1; JOINED.
CC EMBL; AJ297551; CAC21449.1; JOINED.
CC EMBL; AJ297552; CAC21449.1; JOINED.
CC EMBL; AJ297553; CAC21449.1; JOINED.
CC EMBL; AJ297554; CAC21449.1; JOINED.
CC EMBL; AJ297555; CAC21449.1; JOINED.
CC EMBL; AJ297556; CAC21449.1; JOINED.
CC EMBL; AJ297557; CAC21449.1; JOINED.
CC EMBL; AJ297558; CAC21449.1; JOINED.
CC EMBL; AJ297559; CAC21449.1; JOINED.
CC EMBL; AJ297560; CAC21449.1; JOINED.
CC PIR; AS4600; AS4600.
CC Genew; HGNC:8976; PIK3CB.
CC MIM; 602925; -.
CC GO; GO:0016303; P:phosphatidylinositol 3-kinase activity; TAS.
CC GO; GO:0000187; P:activation of MAPK; TAS.
CC GO; GO:0006935; P:chemotaxis; TAS.
CC GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
CC GO; GO:0000074; P:regulation of cell cycle; TAS.
CC InterPro; IPR008938; ARM.
CC InterPro; IPR008973; C2_CaLB.
CC InterPro; IPR000403; PI3_PI4_kinase.
CC InterPro; IPR002420; PI3K_C2.
CC InterPro; IPR003113; PI3K_p85B.
CC InterPro; IPR000341; PI3K_ras_bind.
CC InterPro; IPR001263; PI3Ka.
CC Pfam; PF00454; PI3_PI4_kinase; 1.
CC Pfam; PF00792; PI3K_C2; 1.
CC Pfam; PF02192; PI3K_p85B; 1.
CC Pfam; PF00794; PI3K_rbd; 1.
CC Pfam; PF00613; PI3Ka; 1.
CC SMART; SM00142; PI3K_C2; 1.
CC SMART; SM00143; PI3K_p85B; 1.
CC SMART; SM00144; PI3K_rbd; 1.
CC SMART; SM00145; PI3Ka; 1.
CC SMART; SM00146; PI3Kc; 1.
CC PROSITE; PS00915; PI3_4_KINASE_1; 1.
CC PROSITE; PS00916; PI3_4_KINASE_2; 1.
CC PROSITE; PS00290; PI3_4_KINASE_3; 1.
KW Transferase; Kinase; Multigene family.

FT DOMAIN 800 1050 PI3K/PI4K.
SQ SEQUENCE 1070 AA; 122762 MW; 81135FE93452C00E CRC64;

Query Match 63.2%; Score 36; DB 1; Length 1070;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
:|:|:|:|:|:
Db 245 DYVLQVSGRVEY 256

RESULT 10
P11B_RAT
ID P11B_RAT STANDARD; PRT; 1070 AA.
AC Q9Z1L0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit, beta
DE isoform (EC 2.7.1.153) (PI3-kinase p110 subunit beta) (PtdIns-3-kinase
DE p110) (PI3K) (PI3Kbeta).
GN PIK3CB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Mulder H., Stenson Holst L., Degerman E.;
RT "Phosphatidylinositol-3 kinase and activation of phosphodiesterase 3B
in adipocytes."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PHOSPHORYLATES PTDINS, PTDINS4P AND PTDINS(4,5)P2 WITH A
CC PREFERENCE FOR PTDINS(4,5)P2 (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4,5-
CC bisphosphate = ADP + 1-phosphatidyl-1D-myo-inositol 3,4,5-
CC triphosphate.
CC -!- PATHWAY: Signaling pathways regulating cell growth.
CC -!- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNIT (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AJ012482; CAA10046.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000403; PI3_PI4_kinase.
DR InterPro; IPR002420; PI3K_C2.
DR InterPro; IPR003113; PI3K_p85B.
DR InterPro; IPR000341; PI3K_ras_bind.
DR InterPro; IPR001263; PI3Ka.
DR Pfam; PF00454; PI3_PI4_kinase; 1.
DR Pfam; PF00792; PI3K_C2; 1.
DR Pfam; PF02192; PI3K_p85B; 1.
DR Pfam; PF00794; PI3K_rbd; 1.
DR Pfam; PF00613; PI3Ka; 1.
DR SMART; SM00142; PI3K_C2; 1.
DR SMART; SM00143; PI3K_p85B; 1.
DR SMART; SM00144; PI3K_rbd; 1.
DR SMART; SM00145; PI3Ka; 1.
DR SMART; SM00146; PI3KC; 1.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS50290; PI3_4_KINASE_3; 1.

KW Transferase; Kinase; Multigene family.
FT DOMAIN 800 1050 PI3K/PI4K.
SQ SEQUENCE 1070 AA; 122607 MW; 4E8EB2333E96E4D5 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 1070;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
:|:|:|:|:|:
Db 245 DYVLQVSGRVEY 256

RESULT 11
SYD_PSEAE
ID SYD_PSEAE STANDARD; PRT; 591 AA.
AC Q51422;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (Asprs).
GN ASPs OR PA0963.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
RN [2]
RP SEQUENCE OF 231-591 FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=97136691; PubMed=8982068;
RA Hishida T., Iwasaki H., Ishioka K., Shinagawa H.;
RT "Molecular analysis of the Pseudomonas aeruginosa genes, ruvA, ruvB
RT and ruvC, involved in processing of homologous recombination
RT intermediates."
RL Gene 182:63-70(1996).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE004530; AAG04352.1; -.
DR EMBL; D83138; BAA11815.1; -.
DR PIR; E83524; E83524.
DR PIR; PC4295; PC4295.
DR HSSP; P21889; 1EQR.
DR HAMAP; MF_00044; -; 1.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.

DR InterPro; IPR002312; tRNA-synt_asp.
DR InterPro; IPR004365; tRNA_anti_1.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA_anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMs; TIGR00459; asps_bact; 1.
DR PROSITE; PS50862; AA TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT CONFLICT 588 P -> A (IN REF. 2).
SQ SEQUENCE 591 AA; 66207 MW; 62B278CA0EDE70A4 CRC64;

Query Match 61.4%; Score 35; DB 1; Length 591;
Best Local Similarity 45.5%; Pred. No. 25;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVR 11
|:|:|:|:|:
Db 67 EFVVKITGKVR 77

RESULT 12
SYD_PSEPK STANDARD; PRT; 591 AA.
AC Q88NJ4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, last sequence update)
DT 15-MAR-2004 (Rel. 43, last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (Asprs).
GN ASPS OR PP1213.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016778; AAN6837.1; -.
DR TIGR; PP1213; -.
DR HAMAP; MF 00044; -; 1.
DR InterPro; IPR004115; GAD dom.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_asp.
DR InterPro; IPR004365; tRNA_anti.

DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 1.
DR Pfam; PF01336; tRNA_anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR PROSITE; PS50862; AA TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 591 AA; 66568 MW; B823515D39E52E0F CRC64;

Query Match 61.4%; Score 35; DB 1; Length 591;
Best Local Similarity 45.5%; Pred. No. 25;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVR 11
|:|:|:|:|:
Db 67 EYVQITGKVR 77

RESULT 13
RL31_CAEEL STANDARD; PRT; 122 AA.
ID RL31_CAEEL
AC Q9U332; Q9U331;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE 60S ribosomal protein L31.
GN RPL-31 OR W09C5.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lennard N.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Durbin R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=a;
CC IsoId=Q9U332-1; Sequence=Displayed;
CC Note=No experimental confirmation available;
CC Name=b;
CC IsoId=Q9U332-2; Sequence=VSP 007365;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the L31E family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z82077; CAB63331.1; -.
DR EMBL; Z82077; CAB63332.1; -.
DR WormPep; W09C5.6a; CE20168.
DR WormPep; W09C5.6b; CE20169.
DR InterPro; IPR000054; Ribosomal_L31e.
DR Pfam; PF01198; Ribosomal_L31e; 1.
DR ProDom; PD006030; Ribosomal_L31e; 1.
DR PROSITE; PS01144; RIBOSOMAL_L31E; 1.
KW Ribosomal protein; Alternative splicing.
FT VARSPPLIC 1 52 Missing (in isoform b).
FT /FTId=VSP 007365.
SQ SEQUENCE 122 AA; 14261 MW; 104B26D7EB25CC32 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 122;
Best Local Similarity 54.5%; Pred. No. 7.9;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||:|:|:
Db 20 EYTIHIAHAR 30

RESULT 14
EFTS_CANBF STANDARD; PRT; 269 AA.
AC Q7VRE5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Elongation factor Ts (EF-Ts).
GN TSF OR BFL272.
OS Candidatus Blochmannia floridanus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.
OX NCBI_TaxID=203907;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=22784745; PubMed=12886019;
RA Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candela F.,
RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,
RA van Ham R.C.H.J., Gross R., Moya A.;
RT "The genome sequence of Blochmannia floridanus: comparative analysis
of reduced genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393 (2003).
CC -!- FUNCTION: Associates with the EF-Tu.GDP complex and induces the
exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF-
Tu.GTP complex up to the GTP hydrolysis stage on the ribosome.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the EF-Ts family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----

EMBL; BX248585; CAD83343.1; -.
HAMAP; MF_00050; -; 1.
DR InterPro; IPR001816; EF_TS.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00889; EF_TS; 1.
DR Pfam; PF00627; UBA; 1.
DR TIGRFAMs; TIGR00116; tsf; 1.
DR PROSITE; PS01126; EF_TS_1; 1.
DR PROSITE; PS01127; EF_TS_2; 1.
KW Elongation factor; Protein biosynthesis; Complete proteome.
FT SITE 83 86 INVOLVED IN MG++ ION DISLOCATION FROM EF-
TU (BY SIMILARITY).
FT SEQUENCE 269 AA; 31039 MW; F6C1A31120E96B45 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 269;
Best Local Similarity 58.3%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 12
|||:|:|:
Db 252 EYCIKINFRV 263

RESULT 15
AMPA_AQUAE STANDARD; PRT; 493 AA.
AC O67868;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE (LAP) (Leucyl aminopeptidase).
GN PEPA OR AQ_2099.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358 (1998).
CC -!- FUNCTION: Presumably involved in the processing and regular
turnover of intracellular proteins. Catalyzes the removal of
unsubstituted N-terminal amino acids from various peptides (By
similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
Xbb-, in which Xaa is preferably Leu, but may be other amino acids
including Pro although not Arg or Lys, and Xbb may be Pro.
CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M17.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB000772; AAC07829.1; -.
DR PIR; H70479; H70479.
DR HSSP; P00727; ILAP.
DR MEROPS; M17.UPW; -.
DR HAMAP; MF_00181; -; 1.
DR InterPro; IPR000819; Peptidase_M17_C.
DR InterPro; IPR008283; Peptidase_M17_N.
DR Pfam; PF00883; Peptidase_M17; 1.
DR Pfam; PF02789; Peptidase_M17_N; 1.
DR PRINTS; PR00481; LAMNOPPTDASE.
DR PROSITE; PS00631; CYTOSOL_AP; 1.
KW Hydrolase; Aminopeptidase; Manganese; Complete proteome.
FT METAL 257 257 MANGANESE 2 (BY SIMILARITY).
FT METAL 262 262 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 280 280 MANGANESE 2 (BY SIMILARITY).
FT METAL 339 339 MANGANESE 1 (BY SIMILARITY).
FT METAL 341 341 MANGANESE 1 AND 2 (BY SIMILARITY).
FT ACT_SITE 269 269 POTENTIAL.
FT ACT_SITE 343 343 POTENTIAL.
SQ SEQUENCE 493 AA; 54543 MW; A32B499C7A52065B CRC64;

Query Match 59.6%; Score 34; DB 1; Length 493;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYVIKVSARV 10
||:|:|:|:
Db 484 EYIMKVSSNV 493

Search completed: October 7, 2004, 15:14:01
Job time : 9.18182 secs

THIS PAGE BLANK (USFIC)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 15:14:05 ; Search time 43.3636 Seconds
(without alignments)
87,313 Million cell updates/sec

Title: US-09-336-091-7
Perfect score: 57
Sequence: .. 1 EYVIKVSARVRF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 57 | 100.0 | 316 | 4 Q8WWH6 | Q8wwh6 homo sapien |
| 2 | 40 | 70.2 | 248 | 16 Q8KEI5 | Q8kei5 chlorobium |
| 3 | 39 | 68.4 | 317 | 4 Q14798 | Q14798 homo sapien |
| 4 | 39 | 68.4 | 318 | 4 Q9BUN9 | Q9bun9 homo sapien |
| 5 | 39 | 68.4 | 458 | 16 Q93N48 | Q93n48 coxiella bu |
| 6 | 37 | 64.9 | 323 | 12 Q9IGU9 | Q9igu9 tanapox vir |
| 7 | 37 | 64.9 | 323 | 12 Q9DHN9 | Q9dhn9 yaba-like d |
| 8 | 37 | 64.9 | 357 | 17 O58509 | O58509 pyrococcus |
| 9 | 37 | 64.9 | 591 | 16 Q87Y31 | Q87y31 pseudomonas |
| 10 | 37 | 64.9 | 933 | 5 Q8T4F9 | Q8t4f9 drosophila |
| 11 | 37 | 64.9 | 2072 | 5 Q9VCH1 | Q9vchi drosophila |
| 12 | 36 | 63.2 | 176 | 16 Q82MB7 | Q82mb7 streptomyce |
| 13 | 36 | 63.2 | 259 | 12 Q7TLN3 | Q7tln3 choristoneu |
| 14 | 36 | 63.2 | 409 | 11 Q8CFF1 | Q8cff1 mus musculu |
| 15 | 36 | 63.2 | 590 | 16 Q83BE5 | Q83be5 coxiella bu |
| 16 | 36 | 63.2 | 594 | 16 Q9X156 | Q9x156 thermotoga |

| | | | | | |
|----|----|------|------|-----------|--------------------|
| 17 | 36 | 63.2 | 1064 | 11 Q8BTI9 | Q8bti9 mus musculu |
| 18 | 35 | 61.4 | 146 | 15 Q7ZRN4 | Q7zrn4 chimpanzee |
| 19 | 35 | 61.4 | 175 | 4 Q9NQP3 | Q9nqp3 homo sapien |
| 20 | 35 | 61.4 | 407 | 16 Q98MV2 | Q98mv2 rhizobium 1 |
| 21 | 35 | 61.4 | 437 | 16 Q9CL14 | Q9cll4 pasteurella |
| 22 | 35 | 61.4 | 520 | 16 Q8A569 | Q8a569 bacteroides |
| 23 | 35 | 61.4 | 560 | 12 Q9DWE8 | Q9dwe8 rat cytomeg |
| 24 | 35 | 61.4 | 566 | 16 Q8R624 | Q8r624 fusobacteri |
| 25 | 35 | 61.4 | 591 | 16 Q88NJ4 | Q88nj4 pseudomonas |
| 26 | 35 | 61.4 | 605 | 16 Q8DJS8 | Q8djs8 synechococc |
| 27 | 35 | 61.4 | 933 | 5 Q8SSA9 | Q8ssa9 encephalito |
| 28 | 34 | 59.6 | 74 | 16 Q83HS9 | Q83hs9 tropheryma |
| 29 | 34 | 59.6 | 141 | 13 Q90YV3 | Q90yv3 ictalurus p |
| 30 | 34 | 59.6 | 152 | 17 Q8ZZ01 | Q8zz01 pyrobaculum |
| 31 | 34 | 59.6 | 198 | 16 Q8DBV3 | Q8dbv3 vibrio vuln |
| 32 | 34 | 59.6 | 254 | 2 Q9ALZ3 | Q9alz3 campylobact |
| 33 | 34 | 59.6 | 254 | 16 Q8DBV1 | Q8dbv1 vibrio vuln |
| 34 | 34 | 59.6 | 255 | 16 Q9HUT1 | Q9hut1 pseudomonas |
| 35 | 34 | 59.6 | 264 | 16 Q8F9W3 | Q8f9w3 leptospira |
| 36 | 34 | 59.6 | 269 | 16 Q7VRE5 | Q7vre5 candidatus |
| 37 | 34 | 59.6 | 317 | 16 Q8UDN7 | Q8udn7 agrobacteri |
| 38 | 34 | 59.6 | 322 | 16 Q8EEL8 | Q8eel8 shewanella |
| 39 | 34 | 59.6 | 324 | 12 Q9Q8M9 | Q9q8m9 myxoma viru |
| 40 | 34 | 59.6 | 324 | 12 Q9Q901 | Q9q901 shope fibro |
| 41 | 34 | 59.6 | 409 | 5 Q9NKX4 | Q9nkx4 patinopecte |
| 42 | 34 | 59.6 | 515 | 16 Q81IS3 | Q81is3 bacillus ce |
| 43 | 34 | 59.6 | 556 | 10 Q9SIQ4 | Q9siq4 arabidopsis |
| 44 | 34 | 59.6 | 568 | 10 Q8RYG5 | Q8ryg5 oryza sativ |
| 45 | 34 | 59.6 | 598 | 3 Q875E4 | Q875e4 podospora a |

ALIGNMENTS

RESULT 1

Q8WWH6
ID Q8WWH6 PRELIMINARY; PRT; 316 AA.
AC Q8WWH6;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Tumor antigen MAGE-N.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hepatoma;
RA Sui Y., Ye J., Wu W.;
RT "Cloning of a new gene of MAGE family in human hepatocellular carcinoma."
RT carcinoma."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF443295; AAL37897.1; --
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
SQ SEQUENCE 316 AA; 35409 MW; A463A9A740A089DF CRC64;

Query Match 100.0%; Score 57; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EYVIKVSARVRF 12

Db 288 EYVIKVSARVRF 299

RESULT 2

Q8KEI5
ID Q8KEI5 PRELIMINARY; PRT; 248 AA.
AC Q8KEI5;
DT 01-OCT-2002 (TremBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc protease, putative.
GN CT0704.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514 (2002).
DR EMBL; AE012841; AAM71941.1; -.
DR TIGR; CT0704; -.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002725; DUF45.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01863; DUF45; 1.
DR PROSITE; PS00142; ZINC PROTEASE; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 248 AA; 28214 MW; 45382717E1A8A9F9 CRC64;

Query Match 70.2%; Score 40; DB 16; Length 248;
Best Local Similarity 58.3%; Pred. No. 5.3;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVIKVSARVRF 12
||:|:|:|:|:
Db 14 EYTVKVSQARY 25
RESULT 3
Q14798 PRELIMINARY; PRT; 317 AA.
ID Q14798;
AC Q14798;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MAGE-4 protein.
GN MELANOMA ANTIGEN-4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95369706; PubMed=7642112;
RA Imai Y., Shichiyo S., Yamada A., Katayama T., Yano H., Itoh K.;
RT "Sequence analysis of the MAGE gene family encoding human tumor-
RT rejection antigens.";
RL Gene 160:287-290 (1995).
DR EMBL; D32075; BAA06841.1; -.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
SQ SEQUENCE 317 AA; 35044 MW; 9B9477253FE307C4 CRC64;

Query Match 68.4%; Score 39; DB 4; Length 317;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|:|:|:|:|:
Db 289 EHVVRVNARVR 299
RESULT 4
Q9BUN9 PRELIMINARY; PRT; 318 AA.
ID Q9BUN9;
AC Q9BUN9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Skin antigen, family A, 8 (Melanoma antigen, family A, 8).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor
RT vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002455; AAH02455.1; -.
DR EMBL; BC012744; AAH12744.1; -.
DR EMBL; BT007340; AAP36004.1; -.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
SQ SEQUENCE 318 AA; 35214 MW; EA02C1FB42F6C080 CRC64;

Query Match 68.4%; Score 39; DB 4; Length 318;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|:|:|:|:|:
Db 291 EHVVRVNARVR 301
RESULT 5
Q93N48 PRELIMINARY; PRT; 458 AA.
ID Q93N48;
AC Q93N48;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN CBU0695.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RA Hoover T.A., Vodkin M.H., Williams J.C., Culp D.W., Thompson H.A.;
RT "A chromosomal DNA deletion explains the phenotype of the Coxiella
RT burnetii phase II variant.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=12704232;
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin H., Davidsen T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,

RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coxiella
burnetii.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460 (2003).
DR EMBL; AF387640; AAK71274.1; -.
DR EMBL; AE016962; AAO90239.1; -.
DR TIGR; CBU0695; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.
KW Hypothetical protein; Complete_proteome.
SQ SEQUENCE 458 AA; 53179 MW; A42FF4615563D648 CRC64;

Query Match 68.4%; Score 39; DB 16; Length 458;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
:|:|:|:|:
Db 7 KYLIKIAARIR 17

RESULT 6
Q9IGU9
ID Q9IGU9 PRELIMINARY; PRT; 323 AA.
AC Q9IGU9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE J4R.
OS Tanapox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=99000;
RN [1]
RP SEQUENCE FROM N.A.
RA Essani K., Paulose-Murphy M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF245394; AAF97774.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR004900; Pox_P35.
DR Pfam; PF03213; Pox_P35; 1.
SQ SEQUENCE 323 AA; 37342 MW; A964C9E1DD5880AD CRC64;

Query Match 64.9%; Score 37; DB 12; Length 323;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YVIKVSARVR 12
:|:|:|:|:
Db 187 YIIRVSTAIRF 197

RESULT 7
Q9DHN9
ID Q9DHN9 PRELIMINARY; PRT; 323 AA.
AC Q9DHN9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 74L protein.
GN 74L.
OS Yaba-like disease virus (YLDV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21176366; PubMed=11277691;
RA Lee H.J., Essani K., Smith G.L.;
RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus.";
RL Virology 281:170-192 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of.
DR EMBL; AJ293568; CAC21312.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR004900; Pox_P35.
DR Pfam; PF03213; Pox_P35; 1.
SQ SEQUENCE 323 AA; 37354 MW; 9D1111ABC5E19B27 CRC64;

Query Match 64.9%; Score 37; DB 12; Length 323;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YVIKVSARVR 12
:|:|:|:|:
Db 187 YIIRVSTAIRF 197

RESULT 8
O58509
ID O58509 PRELIMINARY; PRT; 357 AA.
AC O58509;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PH0751.
GN PH0751.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76 (1998).
DR EMBL; AP000003; BAA29842.1; -.
DR PIR; H71122; H71122.
DR InterPro; IPR002831; DUF118.
DR Pfam; PF01978; DUF118; 1.
DR ProDom; PD006327; DUF118; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 357 AA; 40905 MW; 24BE20213B3C402C CRC64;

Query Match 64.9%; Score 37; DB 17; Length 357;
Best Local Similarity 72.7%; Pred. No. 32;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
:|:|:|:|:
Db 297 EYPIEVSGRVR 307

RESULT 9
Q87Y31
ID Q87Y31 PRELIMINARY; PRT; 591 AA.
AC Q87Y31;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aspartyl-tRNA synthetase.
GN ASPS OR PSPTO3981.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA White O., Fraser C., Collmer A.;
RT "Complete sequence of Pseudomonas syringae."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016870; AAC057440.1; -.
DR TIGR; PSPTO3981; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004815; F:aspartate-tRNA ligase activity; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006422; P:aspartyl-tRNA aminoacylation; IEA.
DR InterPro; IPR004115; GAD dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_asp.
DR InterPro; IPR004365; tRNA anti-
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 1.
DR Pfam; PF01336; tRNA anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR PROSITE; PS50862; AA TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 591 AA; 66351 MW; 4A841555305FBF36 CRC64;

Query Match 64.9%; Score 37; DB 16; Length 591;
Best Local Similarity 63.6%; Pred. No. 53;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYVIKVSARVR 11
|||:| :||
Db 67 EYVKVVGKVR 77

RESULT 10
Q8T4F9 PRELIMINARY; PRT; 933 AA.
ID Q8T4F9
AC Q8T4F9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SD03848p.
GN CG10192.
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY084206; AAL89944.1; -.
DR FlyBase; FBgn0039122; CG10192.

DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR003890; IF_eIF4G.
DR InterPro; IPR003891; IF_eIF4G_MA3.
DR Pfam; PF02847; MA3; 1.
DR Pfam; PF02854; MIF4G; 1.
DR SMART; SM00544; MA3; 1.
DR SMART; SM00543; MIF4G; 1.
SQ SEQUENCE 933 AA; 106862 MW; 6ADFDBF51C566543 CRC64;

Query Match 64.9%; Score 37; DB 5; Length 933;
Best Local Similarity 77.8%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 IKVSARVRF 12
|||:| :|||
Db 358 IKISSRVRF 366

RESULT 11
Q9VCH1 PRELIMINARY; PRT; 2072 AA.
ID Q9VCH1
AC Q9VCH1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG10192 protein.
GN CG10192.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,


```
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Goayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Milburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003745; AAF56194.2; -
DR FlyBase; FBgn0039122; CG10192.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR003890; IF_eIF4G.
DR InterPro; IPR003891; IF_eIF4G_MA3.
DR Pfam; PF02847; MA3; 1.
DR Pfam; PF02854; MIF4G; 1.
DR SMART; SM00544; MA3; 1.
DR SMART; SM00543; MIF4G; 1.
SQ SEQUENCE 2072 AA; 226597 MW; A8DEF55D5B3F25D8 CRC64;

Query Match 64.9%; Score 37; DB 5; Length 2072;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 IKVSARVRF 12
Db 1497 IKISSRVRF 1505

RESULT 12
Q82MB7
ID Q82MB7 PRELIMINARY; PRT; 176 AA.
AC Q82MB7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN SAV1743.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
```

```
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005028; BAC69454.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 176 AA; 19152 MW; 6AA9343933A30366 CRC64;

Query Match 63.2%; Score 36; DB 16; Length 176;
Best Local Similarity 70.0%; Pred. No. 26;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VIKVSARVRF 12
Db 115 VLKVAARTRF 124

RESULT 13
Q7TLN3
ID Q7TLN3 PRELIMINARY; PRT; 259 AA.
AC Q7TLN3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Choristoneura fumiferana nuclear polyhedrosis virus (CFMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=208973;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93033705; PubMed=1413988;
RA Lee H.Y., Arif B., Dobos P., Krell P.;
RT "Identification of bent DNA and ARS fragments in the genome of
RT Choristoneura fumiferana nuclear polyhedrosis virus.";
RL Virus Res. 24:249-264(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95297142; PubMed=7778276;
RA Xie W.D., Arif B., Dobos P., Krell P.J.;
RT "Identification and analysis of a putative origin of DNA replication
RT in the Choristoneura fumiferana multinucleocapsid nuclear polyhedrosis
RT virus genome.";
RL Virology 209:409-419(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95297155; PubMed=7778286;
RA Liu J.J., Carstens E.B.;
RT "Identification, localization, transcription, and sequence analysis of
RT the Choristoneura fumiferana nuclear polyhedrosis virus DNA polymerase
RT gene.";
RL Virology 209:538-549(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=96030854; PubMed=7595348;
RA Barrett J.W., Krell P.J., Arif B.M.;
```


SQ SEQUENCE 590 AA; 66754 MW; 33E3AD62742969C8 CRC64;
Query Match 63.2%; Score 36; DB 16; Length 590;
Best Local Similarity 70.0%; Pred.No. 86;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EYVIKVSARV 10
| | | | | : |
Db 66 EYVIKVTGKV 75

Search completed: October 7, 2004, 15:17:01
Job time : 45.3636 secs

THIS PAGE BLANK (USPTO)